PERIODICAL ROOM
GENERAL LIBRARY
UNIV. OF MICH.

VOL. XXIII MARCH 1932 NO. 3

The Journal of Educational Psychology

Devoted Primarily to the Scientific Study of Problems of Learning and Teaching

HAROLD RUGG, *Editor*

CONTENTS

Yearly Variation in the Average Intelligence of Pupils Entering the First Grade. 161
MELVIN ARTHUR STEINER

A Non-intellectual General Factor. 173
HENRY F. ADAMS

Individual Differences in Adults 179
ALONZO G. GRACE

The Measurement of Conflict between Honesty and Group Loyalty 187
J. B. MALLER

A Laboratory Study of the Reading of Familiar Formulas. 192
G. NEVIN REBERT

Timing as a Phase of Skill. 204
COLEMAN R. GRIFFITH

Abstract Intelligence and Art Appreciation. 214
HERBERT A. CARROLL AND ALVIN C. EURICH

Order and Frequency of Observations Recorded by a Group. 221
HENRY BOWERS

The Development of Research in Learning 226
ROBERT A. DAVIS AND C. R. BALLARD

Book Reviews 236

$5.00 per Year · Published Monthly Except June to August by

WARWICK & YORK, INC.

BALTIMORE, MD.

Entered as Second Class Matter Nov. 15, 1921, at the Post Office at Baltimore, Md. under the Act of March 3, 1879; additional entry as second class matter at York, Pa

THE JOURNAL OF EDUCATIONAL PSYCHOLOGY

ESTABLISHED 1910

Editor: HAROLD RUGG

Board of Associate Editors

J. CARLETON BELL ARTHUR I. GATES RUDOLF PINTNER E. L. THORNDIKE
FRANK N. FREEMAN V. A. C. HENMON L. M. TERMAN GOODWIN B. WATSON

The price of the Journal is $5.00 a year in the United States; $5.30 in foreign countries. Subscriptions for less than a year are charged at the rate of 70 cents for each number ordered. Back volumes are $6.00 each; back issues, 90 cents each except when three or more years old, and then $1.00 each.

Subscribers should notify the publishers of change in address at least four weeks before publication of the issue with which the change is to take effect. Claims for non-receipt of an issue will not be honored unless made within two weeks after receipt of the next succeeding number.

Manuscripts for publication, books or other materials for review, and correspondence should be addressed to the publishers.

WARWICK AND YORK, INC., *Baltimore, Maryland*

TEACHERS' HEADQUARTERS

European Plan

Entirely Fire-Proof

Centrally Located

EDWARD DAVIS
Manager

HOTEL RENNERT
BALTIMORE, MARYLAND

The Psychology of Learning
1928 Revision By WILLIAM H. PYLE
$2.20+12¢ postage
Warwick & York, Inc., *Baltimore, Maryland*

The Acid Test

What more exacting test can there be of dictionary leadership than the judgment of the educational world?

Every State that has adopted an unabridged dictionary for use in its schools has chosen exclusively

WEBSTER'S NEW INTERNATIONAL DICTIONARY

The Colleges voted overwhelmingly in favor of Webster as the standard of pronunciation, in answer to questions submitted by the Chicago Woman's Club.

Nearly 100% of all schoolbooks are based on the New International for pronunciation, spelling, compounding, and division of words.

The New International has been universally accepted by those best fitted to judge.

It is in accord with the best modern practice.

Use it in your own schools.

Write for our free booklets of interesting ways to teach the use of the dictionary.

GET THE BEST

G. & C. MERRIAM COMPANY
SPRINGFIELD, MASS.

THE JOURNAL OF
EDUCATIONAL PSYCHOLOGY

Volume XXIII March, 1932 Number 3

YEARLY VARIATION IN THE AVERAGE INTELLIGENCE OF PUPILS ENTERING THE FIRST GRADE

MELVIN ARTHUR STEINER

Supervising Principal of Schools, Ingram, Pennsylvania

Some teachers find fault with the preparation and previous training of each new class as soon as it is promoted to their grade. In some cases the general ability or intelligence of the new grade is criticized and compared unfavorably with preceding classes. Such complaints are generally made to shift responsibility or to challenge the efficiency of some other teacher.

It is true, however, that wise teachers frequently observe differences in the general ability of successive classes or grades from year to year. When several different grades are taught some subject by the same teacher, similar differences in the ability of the various grades are noted. Such differences above the first grade arise from inequalities in educational opportunities and training or from original differences in mental ability. This study is an attempt to test the validity of the latter explanation of these assumed differences in successive grades by studying the IQ's of pupils entering the first grade of a school system over a period of years. The chief purpose is to discover variations in the average intelligence of pupils entering the first grade of a particular school over a ten-year period.

Analysis of Problem

This study involves several specific questions: (1) Is there any regular increase or decrease in the average intelligence quotient of each successive first grade that indicates or suggests a change in the type of community? (2) Is there any significant difference between the average intelligence of any two first grade classes? (3) Is there

any evidence to show that such initial differences in mental ability are reflected in the school records or in the later achievement of these two groups?

Comparison of Average Intelligence Based upon an Individual Test

All pupils entering the first grade of the Ingram School during the four-year period from 1921 to 1924 inclusive were given the Stanford Revision of the Binet-Simon Tests. From these mental ages the IQ of each child was obtained. The following table shows the distribution of these pupils as they entered the first grade according to these IQ's.

Table I.—Distribution of Pupils Entering First Grade from 1921 to 1924 on the Basis of Intelligence Quotients Obtained from the Stanford Revision of the Binet-Simon Tests

Year	1924	1923	1922	1921
IQ	No.	No.	No.	No.
121–125	2	0	0	0
116–120	3	4	0	1
111–115	4	6	1	4
106–110	7	9	5	11
101–105	11	13	7	7
96–100	10	8	16	18
91– 95	5	10	11	11
86– 90	5	2	2	2
81– 85	0	1	2	2
76– 80	1	0	0	0
71– 75	0	1	1	1
Total	48	54	45	57
Average	101.7	101.5	96.8	99.3
SD	8.35	9.25	6.85	8.15
SE average	1.21	1.26	1.02	1.08

The mean IQ for these four years does not vary greatly from the normal. The means for 1921 and 1922 are both below 100 and the other two means are slightly above 100. The average of these four means is 99.8, which indicates a normal group of boys and girls. There is no evidence of regular increases or decreases in the yearly

average IQ during this period, for the lowest average, 96.8, belongs to 1922 and the highest, 101.7, to 1924.

When the difference between the highest and the lowest average is considered, the result is 4.9 ± 1.58 with a critical ratio of 3.10. This difference between the average IQ is mathematically significant, because it exceeds three times its standard error. It is evident, therefore, that the average intelligence of the pupils entering the first grade in 1924 is significantly higher than for those entering in 1922.

Comparison of Average Intelligence Based upon Group Test

With the exception of 1928, every class entering the first grade from 1925 to 1930 inclusive was given the Detroit First Grade Intelligence Test. In 1928 the Pintner-Cunningham Primary Mental Test was used. Although those IQ's are not compared with the others, they are given in Table II to show how the pupils are distributed on the basis of that test. The IQ's for the other five years shown in Table II are all obtained from the Detroit Test.

The average IQ's of first grades over a six-year period is given in the preceding table. Since a different test is reported for 1928, those results are not compared with the other five years. An examination of the averages for the five years reveals no uniform trend toward higher or lower group averages. Since the average for these five years is 106.7, it is evident that this test yields higher IQ's than the test used from 1921 to 1924. When the sizes of the standard deviations shown in Tables I and II are compared, it is seen that the Detroit or group test shows greater variation in individual scores than the Stanford Revision of the Binet-Simon. The range in IQ's for Table I is from 71 to 122 and for Table II from 73 to 141.

When the highest average, 111.8, for 1927 is compared with the lowest, 103.0 for 1925, the difference is 8.8 ± 3.04 with a critical ration of 2.90. This is a significant difference, for it is almost three times its standard error. Although this difference is almost twice as large, as the difference found by comparing the average on the individual test, the critical ratio for the group test is slightly smaller. This is caused by the larger standard deviations of the group test, which indicates a wider distribution of scores on that test.

When the scores on the Pintner-Cunningham Test for the first grade entering in 1928 are examined, it is evident that they compare more favorably with the Stanford Revision of the Binet-Simon scores than with the scores on the Detroit test. The average on this test

is 4.2 points below the average of 106.7 for the five years based upon the Detroit scores and is only 2.4 points above the average of the four years based upon the individual test. No evidence of the reliability of any of these tests is available from this study, but the different groups of pupils tested rank highest on the Detroit, next on the Pintner-Cunningham, and lowest on the Stanford Revision of the Binet-Simon.

TABLE II.—DISTRIBUTION OF PUPILS ENTERING FIRST GRADE 1925 TO 1930 ON BASIS OF INTELLIGENCE QUOTIENTS OBTAINED FROM DETROIT FIRST GRADE TEST

Year	1930	1929	1927	1926	1925	1928[1]
IQ	No.	No.	No.	No.	No.	No.
140–144	0	0	1	0	.....	0
135–139	1	0	2	0	.....	0
130–134	1	0	4	0	.....	0
125–129	3	1	5	2	2	1
120–124	4	10	3	3	0	1
115–119	7	7	2	3	1	5
110–114	5	9	4	2	8	7
105–109	8	2	6	10	7	5
100–104	7	3	5	8	4	1
95– 99	3	5	5	10	6	1
90– 94	3	4	1	3	4	8
85– 89	3	6	4	4	3	1
80– 84	4	0	1	2	1	2
75– 79	1	0	0	0	0	1
70– 74	0	0	0	0	1	2
Total.....	50	47	43	47	37	35
Average.....	106.9	108.4	111.8	103.5	103.0	102.2
SD.....	14.15	12.55	15.85	11.10	11.25	14.40
SE average.....	2.00	1.83	2.42	1.62	1.85	2.44

[1] Pintner-Cunningham Test was used.

COMPARISON OF SCHOOL RECORDS

There are several methods of testing the validity of the IQ's obtained when pupils enter the first grade. In this section an analysis of the school progress of the two pairs of first-grade classes showing significant differences in intelligence is discussed. The individual test given four successive years to all pupils entering the first grade

revealed significant differences between the average IQ's of the classes of 1922 and 1924, and the group test used five years showed that the classes of 1925 and 1927 differed considerably in average intelligence. In the first case the record of each pupil's progress for twelve and sixteen semesters respectively is available. In the second pair of first grades, the progress of each pupil for ten and six semesters is possible. This includes enough semesters of schooling to discover any significant differences in the progress of these four first-grade groups. Table III presents these data by giving the total number of pupil-semesters of attendance, the total number of semesters gained and lost, and the per cent of the total pupil-semesters gained or lost by each of these four first grades.

TABLE III.—COMPARISON OF PROGRESS BY SEMESTERS OF PUPILS ENTERING THE FIRST GRADE IN FOUR DIFFERENT YEARS

Year	Total semesters of attendance	Number of semesters		Difference	Per cent of gain or loss
		Gained	Lost		
Individual Test					
1924	424	14	11	3	0.71
1922	567	12	23	−11	−1.94
Group Test					
1927	228	7	5	2	0.88
1925	317	9	9	0	0

An examination of the preceding table shows that the classes of 1922 and 1925, which have the lowest average intelligence on the individual and the group test respectively, do not have nearly as good a record of semester promotions as the other two classes. In the case of the classes of 1924 and 1927, which rank highest on the individual and the group tests respectively, there is a gain in semesters for each class. It is evident, therefore, that the differences in average intelligence indicated by the two tests are significant enough to show similar differences in pupil progress through the grades.

When the two classes ranking lowest in average intelligence are compared with each other, the class of 1922 seems to have a much lower

record of semester promotions than the class of 1925. Since this study includes no data later than September, 1930, the class of 1922 has school records for sixteen semesters compared to ten semesters for the class of 1924. This difference in length of school life is a partial explanation for a zero result for 1925 compared to a loss of 1.94 per cent of total semesters in 1922 as shown in Table III. When the records, however, of the class of 1922 for the first ten semesters up to the fall of 1927 are compared with the results for the entire sixteen semesters, there is practically no change in the per cent of difference between the number of semesters lost and gained; but poor attendance in the first grade accounts for most of that early loss.

Another possible explanation is the difference in the average intelligence of these two groups. Table I shows that the class entering in 1922 had an average of 96.8 on the Stanford-Binet Test, and Table II shows that the class of 1925 had an average IQ of 103 on the Detroit Test. Although the results of the individual and the group test mentioned are not comparable, it may be safe to assume that the averages for these two first grades indicate a significant difference in ability. Since this is the only other possible explanation for so great a difference in school progress, there must be some agreement between ability as measured by these tests and pupil progress through school.

In comparing the school records of the two classes ranking highest in the individual and the group test respectively, Table III clearly indicates very little difference. The per cents of gain over loss in semesters vary only .17 per cent. The number of semesters gained by the class entering in 1924 is exactly twice as large as that for the class of 1927, and the number of semesters lost by each group shows almost the same relationship. Since the one class had twelve semesters of schooling and the other had only six, this seems to prove that the progress of pupils is almost uniform in the different grades of this particular school. There is considerable difference in the average IQ's of these two classes, for the class of 1924 had an average of 101.7 on the individual test and the class of 1927 had 111.8 on the group test. Although this indicates a difference of 10.1 points between the averages on the two tests, both classes have an average IQ considerably higher than the average of all the classes on each of these tests. The class using the individual test is 1.9 points above the average of the four classes on that test, and the one using the group test is 5.1 points above the average for those five classes. These differences are proportionally

not much greater than the ranges of 4.9 and 8.8 points for the means of the individual and the group test respectively. The slight increase of .17 per cent is net semester-gains for the year 1927 over 1924 is only a tendency in the same direction as the average ability of the two groups and cannot be considered proportional to it.

This analysis of the data shown in Table III seems to indicate an agreement between the relative average ability of classes entering the first grade and their later progress through each successive grade. Although no attempt is made to evaluate these differences, it is clear that they all point in the same direction. When the two pairs of first grades with different averages on the same intelligence tests are compared, the group with the higher average has the larger number of semester gains and the smaller number of losses in each case. Even in the case of the average IQ based upon a group and an individual test, corresponding relationships exist.

Another form of school records that can be used to compare two first grades with significant differences in average initial ability is the final average of all subject marks for each semester. Since the letters A, B, C, D, and E are used to indicate the semester averages of all marks in the various subjects, they are given numerical values of 5, 4, 3, 2, and 1 respectively for statistical purposes. No great claim for the validity of these semester marks is made, but this evidence is presented to supplement the other data.

The averages of these semester marks for the classes entering in 1922 and 1924 respectively are 4.177 and 4.183. In the case of these two groups previously compared, there is practically no difference in their averages. The other two classes, 1924 and 1927, have averages of 4.193 and 4.338 respectively. These have a difference large enough to be somewhat significant. It is very surprising, however, to note that these four averages are all above the mark of B instead of being near C, the middle mark of the scale. This is a possible explanation for the close agreement between the average of the two pairs of first grades considered.

One explanation for this close agreement in the averages of the semester marks for these four classes is that special subjects are included with the regular subjects. Although the marks in the special subjects are not given the same weight in averaging the marks as the subjects to which more time is assigned, some teachers may not have been very careful to observe this principle. The means of the semester marks vary so slightly that the differences are not large enough to be

significant statistically, but both favor the classes having the higher average IQ's.

Comparison of Standard Test Scores

A better basis for the comparison of these first grades is obtained by the use of standard achievement tests. It is fortunate that such records for the pupils of the classes of 1925 and 1927 are available. In November of 1928, the pupils who entered the first grade in 1925 were given Battery A of the Public School Achievement Tests, Form I. The scores of twenty-seven of the thirty-seven pupils that entered the first grade in September, 1925 are comparable with the scores of thirty pupils of the forty-three entering in 1927 on the same test given in November 1930. These two groups of pupils, who remained in the same school for three years under the same teachers except one, and under the same principal, were given a battery of five tests at the beginning of their fourth school year by the writer.

To be able to compare these two groups of pupils remaining from the classes of 1925 and 1927 for the purpose of discovering whether the class of 1927, which had a higher average IQ in the first grade than the class of 1925, has the higher average score on standard tests at the beginning of the fourth year of school, they must be representative samples of each first grade. This is indicated by comparing the average IQ of each sample with the average of the entire first grade. Table II shows that the averages of the class of 1927 and 1925 are 111.8 ± 2.4 and 103.0 ± 1.8 respectively. These averages for those pupils of each class entering the fourth year of school are 111.9 ± 2.8 and 102.9 ± 2.4 respectively. The averages, therefore, of the IQ's obtained from the test used in the first grade, indicate the same relative difference between the samples having standard test scores as between the original two first grades.

Battery A of the Public School Achievement tests consists of the following separate tests: Reading, arithmetic computation, arithmetic reasoning, language usage, and spelling. Only the total scores are used, but similar differences exist for each of the tests. Since these pupils at the beginning of the fourth year of school are found in Grades IIIA, IVB, and IVA, the distribution for each group shows great variability. The range for the thirty pupils that entered the first grade in 1927 is from 55 to 195 and the range for the other group of twenty-seven pupils is from 85 to 205. Such a wide distribution

for small groups indicates large standard errors for their averages. These averages are 149.7 ± 5.4 for the class of 1927 and 145.7 ± 6.1 for the class of 1925. The class of 1927 has an average of four points more than the other class, which is about one-half the standard error of the difference. Although this difference is not mathematically significant, it represents about one-sixth or seventeen per cent of the difference between the norm of 110 for IIIA (high third grade) and that of 134 for IVB. Since the averages of these two groups show a difference of one-sixth of a semester's schooling in achievement at the beginning of the fourth school year, it seems safe to believe that the pupils entering the first grade in 1927 made somewhat greater progress than the other class, which had a lower initial average IQ.

Another method of checking the valadity of the IQ's obtained by testing first grade pupils or of comparing the relative effect of differences in the grades upon their future performance is based upon correlations. Although there is a significant difference between the average intelligence of the first grades entering in 1925 and 1927, the product-moment correlations between each pupil's IQ upon entering school and his score on the achievement test used three years later for the representative sample of the classes of 1925 and 1927 are .62 ± .08 and 64 ± .07 respectively. These correlations are significant and indicate a definite relationship between IQ's and achievement as measured by standard tests three years later. It is not claimed that these relationships are close enough to attempt to predict an individual's performance over a period of three years upon the basis of the group intelligence test used in this study.

These correlations, however, are large enough to prove that there is considerable agreement between the initial ability of first grade pupils and their achievement or attainment three years later. Whether this same relationship exists at the end of a longer period of schooling is an interesting problem but not a definite part of this investigation. It is doubtlessly true that achievement depends upon a number of other factors besides intelligence, such as desire to work, condition of health, social and economic status of the home, and others; but it seems probable that a better index of general intelligence is obtained by testing pupils before they have received any formal schooling than after they have been in school for a few years and have been changed by such training that it becomes increasingly difficult to measure only native ability and not acquired skills or learning.

Summary and Conclusions

In this study, the answers to three specific questions are stated. Besides these answers, several other significant facts are included. These, however, are really supplementary to the chief purpose of this investigation.

It is doubtless true that the average mental ability of pupils entering the first grade of a school system in a community of three or four thousand inhabitants may undergo considerable change within a period of ten years. Such a change, for instance, takes place in a community whenever the type or class of residents begins to be modified by the introduction of new industries. In the present study of the average intelligence of first grade pupils in a residential suburb of Pittsburgh, there is no evidence of a general deterioration or improvement in the average mental ability of children entering the first grade over a ten-year period, but there is considerable variation in this average for each first grade class during this period.

This leads to the evidence of differences between the average mental ability of pupils entering the first grade during each successive year. These differences in the average IQ of successive first grades are based upon the Stanford Revision of the Binet-Simon Intelligence test from 1921 to 1924 inclusive, and upon the Detroit First Grade Intelligence test from 1925 to 1930 inclusive with the exception of 1928 when the Pintner-Cunningham Primary Intelligence test was used. The difference between the highest and the lowest average IQ of the first grades for the years included by the individual test and by the group test respectively is significant enough to believe that the quality of pupils entering the first grade is not uniform from year to year in this or in any similar community.

No explanation for this yearly variation in the quality of childhood presented to the public school of any small community is obtained from the data of this investigation. It is evident that differences occur. There are two possible reasons for such fluctuations in average mental ability: (1) It is a natural condition of human life; (2) it is a normal condition found in studying relatively small groups of pupils.

The first grade classes having the highest and the lowest average IQ on the individual and on the group intelligence test respectively are compared by using school records consisting of semester progress and final term marks and by using standard test scores. This is done to discover the validity of the IQ's obtained from first grade pupils

and the relative permanecy of initial differences in the average mental ability of first grade classes.

Both forms of school records favor the first grade classes with the highest average on the individual and the group intelligence test respectively. These two classes entering in 1924 and 1927 gained .71 and .88 per cent of the total semesters in attendance more than they lost, but the other two classes that are compared with the former show zero gains and a net loss of 1.94 per cent of the total semesters attended. In the case of the averages of all the final semester marks, a slight similar agreement is indicated. Although these differences in school progress and school marks are not mathematically significant, the evidence is so consistent in each case that initial difference in the average ability of these first grades are definitely related to their marks and progress through school.

A final comparison of the classes of 1925 and 1927, which were given the group intelligence test, is based upon the standard test scores of a representative sample of each class tested at the beginning of their fourth year of school. An average difference of four points on the five tests of the Public School Achievement tests favors the class of 1927, which had the highest average IQ of all the first grades taking the group intelligence test. Wide variation in the distribution of the scores of each class and the size of each sample gives such large standard errors, that this difference is only one-half the size of its standard error. Since this difference of four points in the averages of these classes represents about one-sixth of a semester in achievement as measured by this battery of tests, it seems safe to state that the class with the higher average intelligence in the first grade is superior to the other on the basis of the achievement tests taken three years later.

Besides the comparison of the averages of the two samples of the classes entering in 1925 and 1927, the product-moment correlations of .62 and .64 for these two groups respectively between the IQ's obtained from the first grade tests and the achievement test scores made at the beginning of their fourth year of school indicate almost the same degree of relationship for each class. Although these correlations are not large enough to claim any special validity for the intelligence test used, they do indicate considerable relationship between these two measures of mental ability and of achievement after an interval of three years. Two conclusions are natural: (1) The IQ's of first grade pupils based upon a group test are positively related to their scores on

achievement tests three years later, (2) the performance of first grade pupils on a group intelligence test is a partial indication of their success in the higher grades.

This study clearly proves that first grade classes vary sufficiently in mental ability to differ in achievement through several, if not all grades. It is not necessary, however, for any teacher to become discouraged with the quality or the attainment of those pupils coming to him, because it is possible to overcome not only the effects of poor training but even some of the inherent mental weaknesses revealed by an intelligence test. An examination of the individual records used in this study clearly shows that some faithful and skillful teacher was able in many cases to inspire a pupil with low IQ in the first grade to sufficient effort or to diagnose such a pupil's peculiar difficulties in such a way that his later achievement became much better than indicated by his IQ. Although it is true that successive first grades vary significantly in mental ability, their later progress and achievement are only related and not directly proportional to the initial differences indicated by an intelligence test. This condition is the result (1) of the inaccuracy of IQ's, (2) of other essential qualities not measured and, (3) of a superior type of training.

In this study, the value of giving first grades an intelligence test is not definitely mentioned, but it is clearly implied. The fact that some relationship between such IQ's and the achievement three years later exists suggests the use of those IQ's as a partial basis for classification in some of the other primary grades besides the first. The first grade teacher, who uses the results of these intelligence tests first, finds them very helpful in dividing each first grade into three ability groups and in studying each individual pupil. Beyond the first grade, the analysis of individual difficulties or of problem cases always includes the study of those early intelligence test scores.

The writer recognizes the fact that this investigation is incomplete without a comparison of the IQ's obtained in the first grade with those based upon tests given in some higher grade. Since this has been done in other studies and does not seem directly related to the present problem, it has not been considered sufficiently important to collect such additional data by the administration of some other intelligence tests. The real contribution of this study is that some relationship between the IQ's of first grade pupils and their later school progress and achievement exists.

A NON-INTELLECTUAL GENERAL FACTOR

HENRY F. ADAMS

University of Michigan

That Spearman, in the development of his theory of "*g*" and "*s*," has discovered a general principle of correlations rather than one limited to general mental ability, has been maintained by the writer in recent articles.[1] In these, it has been shown that such widely separated phenomena as rainfall and shooting at a target conformed, when tables of correlation coefficients were obtained, to Spearman's tetrad difference criterion.

In the present paper will be presented supplementary evidence in favor of our contention. The material to be discussed consists of the scores made by the most outstanding amateur and professional golfers in the three major tournaments held in this country during the past eleven years. These tournaments are the National Open, the National Amateur, and the one held by the Professional Golfers' Association. In reporting these tournaments, the news-papers furnish the two most essential kinds of information. The first of these is the hole-by-hole scores made by the most skillful and the most popular participants in each of these events. The second is a more or less complete description of the course over which each tournament is played, including the length of each hole. The data to be presented are derived entirely from these sources. Whatever selection of players there may be is therefore made by the sporting writer or editor and not by the writer of this paper. If the hole by hole scores of all qualifiers in the twenty-eight tournaments to be discussed were given, if there were sixty-four qualifiers in each tournament and each played seventy-two holes, the total number of scores would be 129,024. This paper is based upon a modest quarter, and presumably the best quarter, of these.

Since some of the newspapers give the length of each hole of the course, and also the number of strokes each of the selected players takes on each hole, simple mathematical computation furnishes the average number of strokes for each length of hole. Since, furthermore, the different holes of the same course differ one from another in length, it is possible to arrange the different holes by a rank difference method in order of length. Similarly, the different holes can be

[1] *Journal of Applied Psychology*, Vol. XV, 1931, pp. 16–34, 358–377.

arranged in order of merit in terms of the average number of strokes taken to play the hole. Two orders, representing two characteristics of the same hole, are thus obtained. A correlation coefficient will consequently indicate the relationship between the two characteristics.

Such correlation coefficients were computed for the twenty-nine courses on which the twenty-eight tournaments were played. They range in value from .81 to .99 with an average value of .93. Similar figures, obtained from the records of twenty-five minor tournaments, range from .79 to 1.00 with an average of .94. It would appear, consequently, that length of hole is the most important single factor in determining score, for the most skillful group of golfers at least.

Distance is a general or common factor characteristic of all holes of all courses. All holes must have some length, they would not be golf holes unless they did. The usual range is from one hundred to six hundred twenty yards, with an average length of about three hundred seventy-five. Usually no two holes on the same course are of precisely the same length. As a common factor, length is quite similar to Spearman's "*g*."

Closely akin to the specific factors of Spearman, which are uncorrelated and hence vary at random, are the other incidentals of the terrain. Slope of ground would be specific, for the direction and amount of slope would not be the same for any two holes except by chance. Direction of wind would vary from day to day and would, in the long run, be uncorrelated with direction of slope. Similarly, rain, temperature, width of fairway, size and undulation of putting surface, frequency and severity of hazards, to mention only a few, would, as specific factors, be independent variables.

A more beautiful set up for testing the Spearman theory of general mental energy with data in which intelligence is not a factor would be hard to find. Lengths of holes may be correlated with average number of strokes per hole, the result, r_{sl}, being comparable to his r_{ag}. The orders representing the average number of strokes per hole in relation to length for a certain course may be correlated with the similar order for any other course. Thus a correlation table may be made to which the tetrad difference, or its equivalent, may be applied. Should it be found to hold, a general factor of distance will have been discovered.

Three such tables will be presented. This is done in part to guard against any possible charge of coincidence. A more serious reason, however, is to be found in the fact that each expresses a somewhat different aspect of the game of golf. In the National Open the partici-

pants include both amateurs and professionals. Medal play is the rule and the winner is the one who makes the four rounds in the smallest number of strokes. In both the National Amateur and the P. G. A. tournaments, match play or the winning of individual holes is the custom. In each match one participant is eliminated. Better golf is supposed to be played in medal rather than in match play.

In Table I are presented the correlations for the National Open tournaments for the years 1921–1931 inclusive. Table II gives the

TABLE I.—NATIONAL OPEN 1921–1931 INCLUSIVE

Course	Col.	Sko.	Inw.	O. H.	Wor.	Sci.	Oak.	Oly.	W. F.	Int.	Inv.
Columbia	(86)	85	83	76	83	81	83	82	83	83	82
Skokie	89	(99)	96	88	96	93	96	95	96	96	95
Inwood	83	96	(97)	86	94	91	94	93	94	94	93
Oakland Hills	76	88	86	(89)	86	84	86	85	86	86	85
Worcester	80	96	94	88	(97)	91	94	93	94	94	93
Scioto	90	93	92	87	88	(94)	91	90	91	91	90
Oakmont	86	97	96	89	96	93	(97)	93	94	94	93
Olympia	79	95	92	82	93	86	90	(96)	93	93	92
Winged Foot	83	96	91	92	95	91	94	93	(97)	94	93
Interlachen	76	93	92	91	96	89	93	91	94	(97)	93
Inverness	79	94	97	88	93	88	94	88	90	93	(96)

Average r_{ag} = .950
Average r^2_{ag} = .903
Average r_{ab} = .907
Average discrepancy = .02 ± .03.

TABLE II.—NATIONAL AMATEUR 1922–1931 INCLUSIVE

Course	Bro.	Flo.	Mer.	Oak.	Bal.	Min.	B. B.	P. B.	Mer.	B. H.
Brookline	(96)	90	85	91	84	93	88	92	84	91
Flossmoor	88	(94)	84	89	83	91	87	90	82	89
Merion	85	89	(89)	85	78	86	82	85	77	85
Oakmont	89	92	86	(95)	84	92	87	91	83	90
Baltusrol	83	79	88	81	(88)	85	81	84	77	84
Minikahda	96	90	87	93	86	(97)	89	93	84	92
Brae Burn	87	92	85	88	76	91	(92)	88	80	78
Pebble Beach	90	91	86	91	91	92	88	(96)	84	91
Merion	82	83	86	90	76	85	80	80	(87)	83
Beverley Hills	92	90	84	93	82	90	86	90	81	(95)

Average values of r_{ag} = .929
Average value of r^2_{ag} = .863
Average value of r_{ab} = .869
Average discrepancy = .02 ± .04.

corresponding correlation coefficients for the National Amateur tournaments for the years 1922–1931 inclusive. Table III presents the coefficients for the P. G. A. tournaments for seven years, 1924–1931, with 1930 omitted.

TABLE III.—P. G. A. TOURNAMENTS FOR 1924, 1925, 1926, 1927, 1928, 1929, AND 1931

Course	F. L.	Oly.	Oly.	Sal.	C. C.	F. F.	Hil.	Wan.
French Lick	(87)	70	79	80	83	77	78	81
Olympia #3	88	(81)	74	75	77	72	73	75
Olympia #4	80	84	(91)	84	87	81	82	85
Salisbury #4	80	69	85	(92)	87	82	83	86
Cedar Crest	81	77	88	92	(95)	85	86	88
Five Farms	85	81	78	71	80	(89)	80	83
Hillcrest	79	75	85	82	83	78	(90)	84
Wanamoisett	78	72	79	82	87	85	74	(93)

Average value of r_{ag} = .898
Average value of r^2_{ag} = .806
Average value of r_{ab} = .806
Average discrepancy = .04 ± .05.

In each of the three tables three series of values are given. Running down the diagonal of each table and enclosed in parentheses are the equivalent of the Spearman r_{ag} values. In the lower left portion appear the empirically determined r_{ab}, r_{ac}, r_{bc}, etc. values. In the upper right section are given the corresponding values determined by the use of the formula $r_{ab} = r_{ag} \cdot r_{bg}$, etc. The following relations have been found to hold.

Tournament	r_{ag}	r^2_{ag}	r_{ab}	Discrepancy
National Open	.950	.903	.907	.02 ± .03
National Amateur	.929	.863	.869	.02 ± .04
P. G. A.	.898	.806	.806	.04 ± .05

These figures conform well with the demands of the Spearman theory. The average value of all the products of the numbers running down the diagonal is .903 for the National Open, .863 for the Amateur and .806 for the Professional tournaments. These are the values demanded by theory. The empirically obtained values are .907, .869, and .806 respectively. The average discrepancy between the

two sets of values runs from ½ to ⅘ of the probable error; hence the agreement could scarcely be better.

This series of relations also re-emphasizes the fact that the r_{ab}, r_{ac}, r_{bc}, etc. correlation coefficients are the *products* obtained by multiplying the r_{ag}, r_{bg}, r_{cg}, etc. values one by the other. The criterion of the tetrad difference depends upon the fact that the coefficients are products, for on this assumption $r_{ac} \cdot r_{bd} - r_{ad} \cdot r_{bc} = r_{ag} \cdot r_{cg} \cdot r_{bg} \cdot r_{dg} - r_{ag} \cdot r_{dg} \cdot r_{bg} \cdot r_{cg} = 0.00$, for the terms on both sides of the minus sign are identical and consequently cancel. If any one of the empirically obtained coefficients is not equal to the theoretical product, the tetrad difference becomes a real number, and if the departure is too great, the criterion is no longer satisfied. It would appear, then, that the tetrad difference is a long, tedious, and complicated method for determining the same fact that is brought out by comparing the empirically and theoretically obtained coefficients and noting the discrepancy or difference between them. The latter method has the saving graces of brevity and simplicity.

Correlation coefficients can be products of two numbers only under one set of conditions, however. For it will be noted in all cases where the tetrad difference criterion holds, no matter whether it be in intelligence tests, rainfall, target shooting, or golf, that the same fundamental variant is being measured in all of the inter-correlated tests. There is some cause for each of the reactions and each reaction tends to approach more or less closely a certain known or unknown standard. When the standard is known, the r_{ag} values can be computed directly; when unknown they can be arrived at by the algebraic treatment of the correlation coefficients. But a standard single either in form or in content is one of the necessary conditions for the working of the tetrad difference criterion and all that it implies.

A second necessary condition is the presence of only uncorrelated errors, for when correlated errors are present they are revealed by the fact that r_{ab} is greater than $r_{ag} \cdot r_{bg}$. The reason for this condition is easy to see if we use the rank difference to illustrate the conditions. We shall assume three ranks each of which approximates closely to the standard order, say the correlation is .90. In the first, the disagreement with the standard occurs in the first third; in the second in the second third; and in the third in the last third. The average intercorrelation of 1 with 2, 1 with 3, and 2 with 3 will tend to be .81, or the square of .90. If, however, combinations of 1 with 1, 2 with 2, or 3 with 3 appear more often than chance would warrant, the correla-

tion will be greater than .81. Since this condition can occur only when positively correlated errors are present, it follows that positively correlated errors reveal their presence when r_{ab} is significantly greater than $r_{ag} \cdot r_{bg}$. By the same line of reasoning, it should follow that if negatively correlated errors are present, their effect should be to make r_{ab} less than $r_{ag} \cdot r_{bg}$.

From the evidence presented the conclusion may be drawn that correlation coefficients are products of two numbers when the standards are identical either in form or in content and when the errors are uncorrelated. This conclusion would imply that the Spearman methods were sound as methods, but that the application of sound methods to a limited range of the total field to which they may apply may lead to awkward and embarassing conclusions and to interpretations too narrow and limited to be true. Whenever there is ability, or capacity, or tendency to approximate to any standard whatever, be it intellectual, or motor, or aesthetic, or humorous, so long as the standard is kept constant throughout, and so long as the errors are uncorrelated, the correlations will conform to the tetrad difference criterion.

INDIVIDUAL DIFFERENCES IN ADULTS

ALONZO G. GRACE

University of Rochester

Variation in adult ability is an important aspect of the educational process. It is the conclusion of this study that should all the average be organized in groups, and all the superior in other groups, the range in accomplishment at the end of a given period would be as great as had the segregation not taken place. In other words, variation within groups is always greater than that between groups. The matter of individual differences, however, must be recognized in the organization of the educational program whether it be in the direction of formal or informal processes. To attempt to mold each individual according to the same educational pattern is an error difficult to rectify and certainly no educational institution is justified in applying this scheme to the masses. Each individual must be fitted according to his needs, interests and abilities. To have done this successfully is the best index of success in the educational process.

There are many factors that influence ability. It seems rather poor educational administration to organize the entire program on the basis of intelligence quotients without partialling out such factors as ability of the instructor to teach, economic background of the student, fatigue, the element of time in testing, interest of the student and a dozen other factors.

The Difference of Age

There apparently is no relationship between mental ability and age, nor is there any evidence to show that age has any influence on achievement. Age, therefore, should not be considered, so far as adult education is concerned, as having much bearing on either the learning process or the amount, rate and speed of intellect. Age, however, is an important factor to be considered in the organization and administration of a program or in the teaching process.

The one fundamental character of age that has been noted is the fact that those who are interested in the formal processes, certificates and diplomas usually present a median age of approximately twenty-one. The median, of course, represents that point in a scale on either side of which one-half the measures fall. Those interested in the more informal types of education are considerably

older. They are usually satisfied with their job. Education to them is a process of self-improvement and not merely so many requirements to be eliminated before the goal is reached. So far as ability is concerned, therefore, age is a factor that may well be disregarded. From an administrative point of view, however, it is not wise to admit children to adult groups, nor is it wise for leaders or instructors to apply to adults a technic of instruction suitable for children.

"In general, nobody under forty-five should restrain himself from trying to learn anything because of a belief or fear that he is too old to be able to learn it. Nor should he use that fear as an excuse for not learning anything which he ought to learn. If he fails in learning it, inability due directly to age will very rarely, if ever, be the reason." Thus did Throndike and his colleagues in 1928 answer William James who in 1893 wrote that "Outside of their own business, the ideas gained by men before they are twenty-five are practically the only ideas they shall have in their lives. They cannot get anything now. Disinterested curiosity is past, the power of assimilation is gone. Whatever individual exceptions might be cited to these are of the sort that prove the rule." Thorndike in his study on adult learning concludes that:

1. Failure to learn in adulthood is due to: (*a*) Lack of capacity to learn a particular thing, (*b*) Lack of strong desire to learn, (*c*) Inadequate ways and means of learning.

2. Provision of educational opportunities for adults is a productive investment for the nation.

3. Age, in itself, is a minor factor in either success or failure. Capacity, interest, energy and time are the essentials.

4. Adult students, by reason of their interest, are not freed from any of the general difficulties encountered by the young student. They have often a motive ulterior to the study itself.

5. Facts of adult learning may encourage industry to face changes in machinery, processes, and the like, with the hope of reducing disturbances by education of the workers.

6. Changes in our general schemes of education, the allotting of proposed additional hours of schooling over adult years instead of extending the age of compulsory attendance to seventeen or eighteen chief factors for this accomplishment: (*a*) Better selection of persons to be taught; (*b*) better selection of content of instruction; (*c*) Better arrangement and sequence of learning—"First the need, then the knowledge or technique to satisfy the need," Dewey; (*d*) Prevention of loss of abilities by forgetting or of time by relearning; (*e*) Lessening of lag of schooling behind science and technology; (*f*) Providing pupils with instruction, and satisfaction of self-respecting accomplishment.

7. Placement of schooling should harmonize with a scheme of (*a*) Productive labor, (*b*) Recreation, (*c*) "Higher life."[1]

[1] Thorndike: "Adult Learning." MacMillan, 1927.

Figure 1 shows the distribution of mental ability of 3819 adults in Cleveland as revealed by the composite results on six tests.

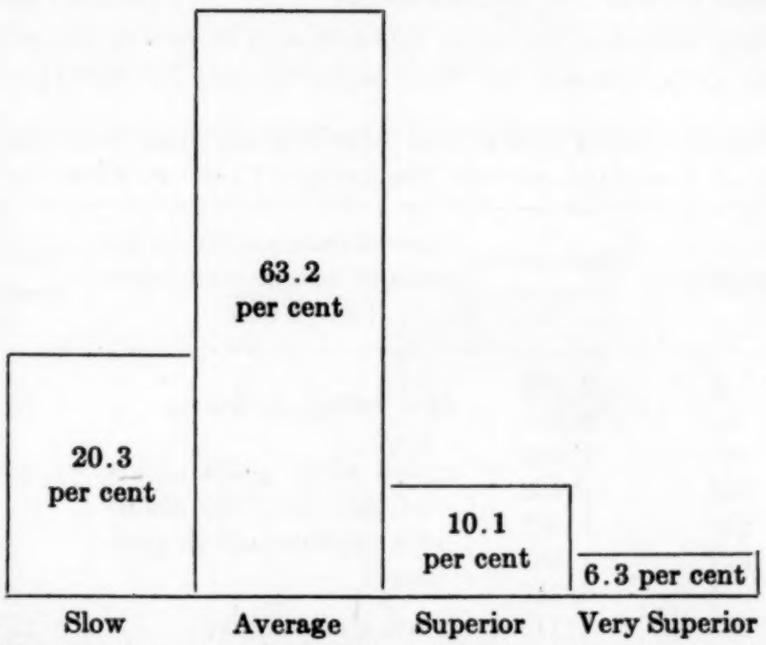


FIG. 1.—The distribution of mental ability of adults as shown by the averager on six tests given to 3819 participants of the educational program for adults sponsored by the Cleveland Board of Education.

These tests indicate that education must come to mean the development of the individual to fit his environment. It shall mean providing the individual with those tools that will enable him tc adjust to new or changing environments with the least possible disturbance to his own sense of social balance and to the equilibrium of society. It is evident, then, that in our educational process, the efforts of educators must be directed more and more toward the influencing of human behavior rather than grinding out certificates, diplomas, and degrees.

The results of Thorndike test are shown in Table I.

The median on this test was 66.7. Sixty-three per cent of the group attained scores that would indicate an ability varying from a capable ninth grader to a capable college graduate. Only 16.6 per cent of the entire group scored below that point on the scale that a capable seventh grader would be expected to attain.

The distribution of ability of 2400 adults as measured by the Thorndike Test is shown graphically in Chart III.

The high score of ninty-eight was obtained by a woman forty-three years of age, American born, and having two years of college training. The low score was zero, or complete failure. This point was reached

eight times. All those who scored from zero to five inclusive, were recent Negro migrants from the south.

There was a negative correlation of −.19 between age and score on the test which would indicate no relationship between age and learning or age and achievement on this test. Table II contains only one

TABLE I.—THE RANGE OF SCORES AND APPROXIMATE ABILITY OF 2400 ADULTS IN CLEVELAND AS REVEALED BY THE THORNDIKE TEST OF WORD KNOWLEDGE

Score	Frequency	Cumulative frequency	Approximate ability on the basis of norms established by the test[1]	Total cases	Per cent total
95–100	40	2,400	Able college graduates	120	5.0
90– 94	80	2,360			
85– 89	138	2,280	Above ninth grade ability but less than the ability of a capable college graduate	1,168	48.7
80– 84	225	2,142			
75– 79	248	1,917			
70– 74	312	1,669			
65– 69	245	1,357			
60– 64	223	1,112	Ninth grade ability	223	9.3
55– 59	194	889	Eighth grade ability	194	8.0
50– 54	185	695	Seventh grade ability	299	12.4
45– 49	114	510			
40– 44	88	396	Sixth grade ability	155	6.5
35– 39	67	308			
30– 34	43	241	Fifth grade ability	88	3.7
25– 29	45	198			
20– 24	37	153	Fourth grade ability	37	1.5
15– 19	28	116	Less than fourth grade ability	116	4.9
10– 14	26	88			
5– 9	23	62			
0– 4	39	39			

[1] Ability based on norms established by the test.

significant fact, *viz.*, that less than two per cent of the group under twenty scored a fourth grade ability or less, whereas one hundred thirty-five or 11.4 per cent of those of voting age were in this group. Age *per se*, however, was a minor factor and of little significance in this instance.

INFLUENCE OF REMOTE ANCESTRY

There was absolutely no relationship between race or nationality and score on this test. There was distinct relationship, however, between score on the test and previous opportunity for education.

The Negro group, mainly southern migrants, scored lowest as would be expected from a study of their previous educational opportunities. This test shows beyond any doubt the need to study carefully the background of an individual before arriving at any preconceived notion of his inherent inferiority. Opportunity for education is a

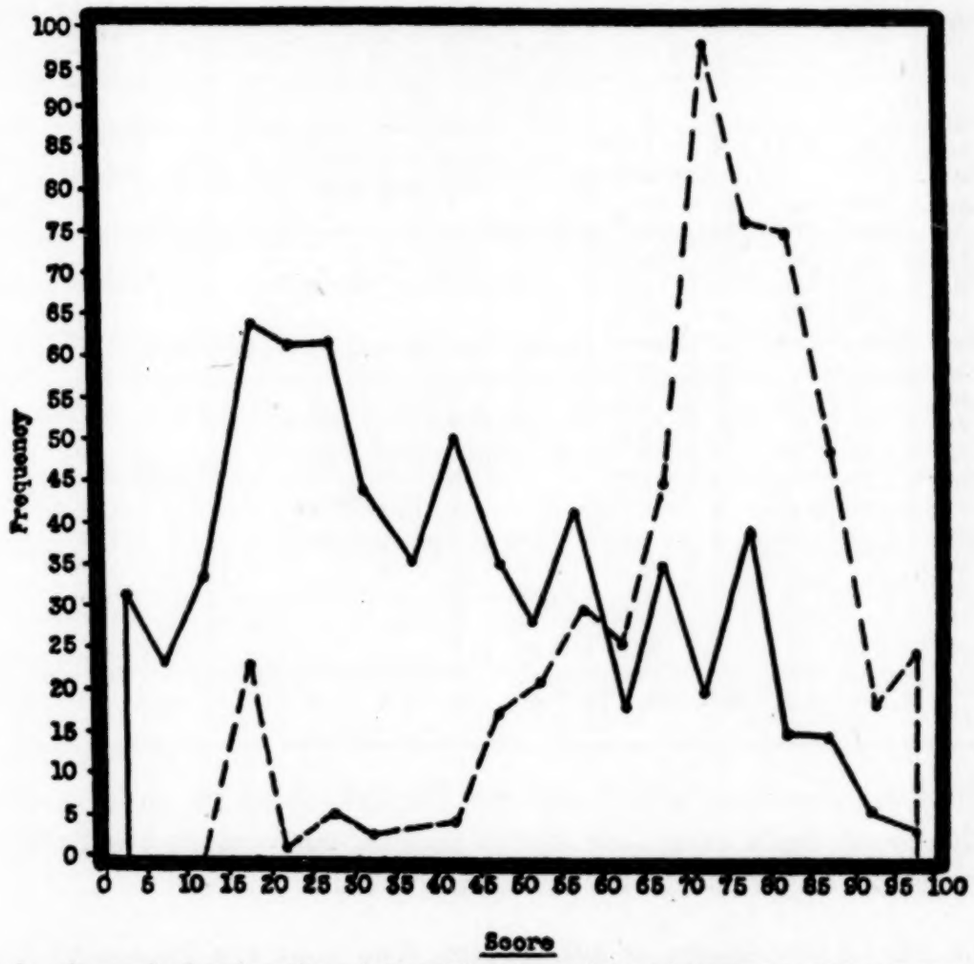


FIG. 2.—A graph showing the relationship of previous education to the scores on the Thorndike Test of Word Knowledge in which 2400 adults participated.
——————One to eight years education
- - - - - - - - - - - - - - - -Ninth grade of high school to two years of college education

factor not to be overlooked in the study of the Negro and the foreign born.

The foreign born or those unable to use the English language with facility would naturally have difficulty in scoring well on a test which involved the use of English—and what test does not? One wonders what would be the result on intelligence quotients, if tests were to be given in the language with which the individual is familiar and involv-

ing the experiences which he daily encounters or with which he was familiar in his home land.

TABLE II.—SCORES OF 2400 ADULTS IN CLEVELAND WHO TOOK THE THORNDIKE TEST OF WORD KNOWLEDGE, BY AGE GROUPS

Scores	Under 20	20–24	25–29	30–34	35–39	40–44	45–49	50–54	55–59	60–64	65–69	Not given	Total
95–100	8	18	5	2	3	..	..	2	..	..	..	2	40
90– 94	22	33	6	5	7	2	..	..	1	..	..	4	80
85– 89	55	49	9	7	5	..	..	2	..	..	..	11	138
80– 84	106	55	21	16	7	2	3	2	..	..	..	13	225
75– 79	109	75	23	18	6	2	3	..	..	..	..	12	248
70– 74	161	80	27	19	7	..	2	..	..	..	..	16	312
65– 69	112	67	39	7	9	..	..	..	..	..	..	11	245
60– 64	105	54	29	11	5	..	..	..	..	..	..	19	223
55– 59	91	44	22	15	5	..	..	..	..	..	..	17	194
50– 54	96	41	24	6	1	..	..	..	..	..	..	17	185
45– 49	50	29	11	4	4	2	1	..	..	..	..	13	114
40– 44	27	23	11	10	2	2	1	..	..	..	..	12	88
35– 39	30	16	7	5	..	..	..	..	..	..	..	9	67
30– 34	17	16	4	1	..	..	1	..	..	..	..	4	43
25– 29	21	7	6	4	..	..	..	1	1	..	..	5	45
20– 24	7	13	4	4	2	2	1	..	..	..	..	3	37
15– 19	7	4	7	3	..	1	2	1	..	..	..	3	28
10– 14	3	2	2	2	4	9	2	..	..	..	..	2	26
5– 9	1	2	1	5	6	5	1	..	..	..	..	2	23
0– 4	2	1	2	2	5	19	8	..	..	..	..	...	39
	1,030	629	260	146	78	47	25	8	2	...	...	175	2,400

One wonders how well a superior English speaking child or adult would do on tests involving Bantu habits, customs and culture, or Chinese life, or the life of the American Indian.

TABLE III.—A COMPARISON OF 2400 ADULTS, WHO TOOK THE THORNDIKE TEST, BY REMOTE ANCESTRY AND WITH RESPECT TO MEAN, MEDIAN, STANDARD DEVIATION, AND PROBABLE ERROR

	Negro	Foreign	Native white
Mean	28.45	56.30	70.05
Median	22.3	63.1	72.1
Standard deviation	24.20	25.25	15.3
PE	1.22	.62	.50

This is further analyzed in the following table. The opportunity of the native white is again evident.

TABLE IV.—COMPARISON OF THE GROUPS WHO TOOK THE THORNDIKE TEST

	Negro and foreign	Negro and white	Foreign born and native white
Difference	27.85	41.60	13.75
$PE_{diff.}$	1.37	1.32	.796
Critical Ratio	20.33	31.52	17.27

THE INFLUENCE OF PREVIOUS TRAINING

The most significant feature of the study was that concerned with previous training of the participants. Table V shows this distribu-

TABLE V.—DISTRIBUTION OF SCORES OF 1216 ADULTS ON THE THORNDIKE TEST OF WORD KNOWLEDGE TABULATED ACCORDING TO PREVIOUS TRAINING OR IN TERMS OF YEARS OF EDUCATION RECEIVED BEFORE ENTERING CLASSES FOR ADULTS IN CLEVELAND

Score	1	2	3	4	5	6	7	8	9	10	11	12	Two years college and over	Total
95–100	...	..	..	..	..	..	...	4	1	11	5	7	10	38
90– 94	...	..	..	..	..	..	...	6	4	4	4	2	7	27
85– 89	...	..	..	..	..	..	2	14	12	11	5	14	6	64
80– 84	...	..	..	1	..	..	...	15	19	16	15	3	27	96
75– 79	...	..	..	1	..	..	3	36	17	19	19	21	7	123
70– 74	...	..	..	1	..	..	...	19	22	19	17	19	27	124
65– 69	...	..	..	1	..	..	4	30	12	14	11	5	11	88
60– 64	...	..	..	7	..	..	...	11	17	8	1	..	..	44
55– 59	...	..	..	..	..	..	14	28	5	19	6	..	..	72
50– 54	1	..	..	..	..	..	11	16	14	5	2	..	..	49
45– 49	...	..	1	..	2	2	12	18	8	8	1	..	..	52
40– 44	4	..	..	9	..	18	14	5	3	1	1	..	..	55
35– 39	2	3	..	11	2	1	13	4	1	...	3	..	..	40
30– 34	1	2	8	6	4	1	16	6	1	...	1	1	..	47
25– 29	12	2	5	12	6	1	19	5	2	1	3	..	..	68
20– 24	10	1	1	28	1	3	10	8	2	...	..	..	..	64
15– 19	28	..	3	12	..	..	11	10	11	...	..	..	..	75
10– 14	25	..	..	..	..	..	7	2	...	...	..	..	..	34
5– 9	21	..	..	..	..	..	1	2	...	...	..	..	..	24
0– 4	32	..	..	..	..	..	...	...	...	...	..	..	..	32
Total	136	8	18	89	15	26	137	239	151	136	94	72	95	1,216

tion for Thorndike study. The great number who received less than a sixth grade education are massed at the low end of the scale. Those receiving more than a sixth-grade education are massed at the upper end of the scale.

Conclusion

1. Every adult group should be given a vocabulary test involving the words most frequently used by an instructor in order that the instructor may either define terms used in a course or simplify his language to meet the vocabulary abilities of his group.

2. Courses for adults must be organized to meet individual differences due primarily to lack of previous educational opportunities. The study indicates that age and remote ancestry have little to do with mental ability.

3. There must develop a guidance program for adults, both for those who desire to continue their education as well as for those who seek opportunities to use knowledge and information. The composite gauge of mental ability of adults shows conclusively that at least twenty per cent of the 3,819 adults studied could not possibly profit from the instruction received.

4. There is no significant relationship between age and mental ability as revealed to the results of intellegence tests. While the results are not shown in this study, correlations between age and mental ability on the base of scores obtained in the Otis, Terman and Ohio State Psychological Test were too small to be of significance. The Thorndike Test revealed a negative correlation.

5. Without doubt much of the educational process is wasted in adult groups because of the deficiency in vocabulary background. For this reason, there probably is a significant relationship between interest in a group and vocabulary span.

THE MEASUREMENT OF CONFLICT BETWEEN HONESTY AND GROUP LOYALTY

J. B. MALLER

Institute of School Experimentation

Teachers College, Columbia University

Considerable progress has been made recently in the measurement of some tangible aspects of character. Tests have been devised which measure with a fair degree of reliability an individual's behavior in situations involving honesty, cooperativeness, inhibition, etc. Whatever these tests measure they register the person's solution to a problem, to a conflict between what is conventionally considered right or wrong. The solution to the problem is either in the direction of honesty or deceit, helpfulness or selfishness, self-control or obedience to impulse.

It was the purpose of this research to study the behavior of children in a controlled situation involving a conflict between group loyalty and honesty. The study was conducted in cooperation with the Character Education Inquiry, which was sponsored by the Institute of Social and Religious Research, and directed by Drs. Hartshorne and May.

The experimental situation consisted of a combination of a cooperation test and an honesty test,* given to two hundred fifteen pupils of Grades VII and VIII. A contest was arranged among eight classrooms of a school "to find out which class can do the best" on a test of simple additions. The children of each class were instructed not to write their names on their papers but the names of their respective groups (such as VIIA or VIIB). The only motive for doing well on this test was to help one's classroom gain recognition. Each test blank, however, had an identifying symbol and the exact amount of work done by each child when working for the group was thus ascertained. The test material included, in addition to the speed test, an honesty test involving the illegitimate use of a key. This test yielded a record indicating the extent to which a child accepted the opportunity for deception in order to help his group.

---

* For a description of these tests see (1) "The Self-marking Test (An Objective Test of Honesty)," J. B. Maller, Bureau of Publications, Teachers College, Columbia University, 1930; (2) "Studies in Service and Self-control," Hartshorne, May and Maller, Macmillan, 1929.

The same test material including the honesty test was given again, but this time the children were told to write their names on their papers. The directions stated that the purpose of the test was "to find out who is the fastest worker in the class, the second fastest, and so on to the very slowest." Each child was thus urged to work as fast as possible for the sake of recognition.

The discrepancy between the speed of work for the group and speed of work for self was then determined for each child. The records of this test also indicated the extent to which a child accepted the opportunity for deception in order to gain personal recognition.

It should be noted that the subjects studied did not know that their work for the group was measured, as no names were to be written on the group work. Nor were they aware of the fact that the illegitimate use of the key, while working for the group or self, was recorded.

The following records were then available for each of the two hundred fifteen children:

1. Degree of cooperation when working for the group on the bona fide speed test.
2. Degree of deception when working for self.
3. Degree of deception when working for the group.

The score of cooperativeness was expressed in terms of the ratio of the speed of work for the group to the speed of work for self.* The two honesty scores were expressed in terms of the percentage of opportunities for deception accepted. Table I gives the means and the standard deviation of these three scores.

TABLE I.—SCORES OF COOPERATION AND HONESTY

Type of score	Mean	Standard deviation	Number of cases
Cooperation	82.2	11.4	215
Deception in work for self	56.4	27.0	215
Deception in work for group	48.0	29.1	215

Table I reveals that the speed of work for group was lower than the speed of work for self, the former being only eighty-two per cent of the latter. The degree of deception when working for self was

* A comprehensive analysis of this measure of cooperativeness will be found in "Cooperation and Competition," J. B. Maller, Bureau of Publications, Teachers College, Columbia Univ., 1929.

significantly greater than the degree of deception when working for the group. The difference between the means was 4.3 times the standard deviation of the difference.

Three groups. The children were divided into three groups, on the basis of their cooperation scores: The most cooperative (those whose scores were more than one SD above the mean), the average (those whose scores fell between one SD above and one SD below the mean), and the least cooperative (those whose scores were more than one SD below the mean). The average deception scores of these groups are shown in Table II.

TABLE II.—DECEPTION SCORES OF THREE GROUPS

Group	Deception in group work	Deception in work for self	N
(*A*) Most cooperative	52.8	54.8	34
(*B*) Average	47.9	56.1	145
(*C*) Least cooperative	38.1	56.7	36

Table II reveals that deception in work for self decreased slightly with increase in cooperativeness, while deception in group work increased consistently with increase in cooperativeness. The children of Group *A*, who were most cooperative, were least honest, while those of Group *C*, the least cooperative, were most honest when they were working for their group. From the point of view of the conflict between honesty and cooperativeness, the children of Group *A* solved the conflict in favor of cooperativeness, while those of Group *C* solved it in the direction of honesty. If both cooperativeness and honesty are considered as learned behavior patterns or as acquired habits, the results would indicate that in Group *A* the habit of cooperativeness was over-developed, at the cost of honesty. In Group *C* the habit of honesty was more developed than that of cooperativeness.

Correlation.—The cooperation score was correlated with each of the two honesty scores, and the latter two scores were correlated with one another. The three inter-correlations were as follows:

Correlation between cooperation and honesty in work for self, $+.16 \pm .04$
Correlation between cooperation and honesty in work for group, $-.28 \pm .04$
Correlation between the two honesty scores, $+.47 \pm .04$

The most significant fact revealed by the inter-correlations was the difference between the first two correlations. The correlation

of −.28 between cooperation and honesty in group work indicates that those who scored high on the cooperation test scored low in honesty. Apparently they tended to increase the score of their group by means of deception. On the other hand, there was a low but positive correlation between cooperation and honesty in work for self.

The correlation of .47 between the two honesty scores indicates that there was some association between honesty in work for self and honesty in work for the group. This is probably the result of a general attitude toward honesty and to the common elements in two testing situations.

The partial correlations throw additional light on the relationship between these variables. When the cooperation score was kept constant the correlation between the two honesty scores rose from .47 to .55. When the honesty in self work was partialled out, the correlation between honesty in group work and cooperation changed from −.28 to −.59. When the score of honesty in group work was kept constant the correlation between honesty in work for self and cooperation changed from +.16 to +.35.

Sex Differences.—No significant differences were found between the general averages of boys and girls on any of the honesty and cooperation scores. An interesting sex difference, however, was found when the results were examined for the separate classrooms. In four out of the six classrooms tested the girls scored somewhat higher than the boys in honesty of group work, while in the remaining two classrooms the girls scored lower in honesty.

Further analysis of the data revealed that in each of the four classes where the girls scored higher in honesty the number of girls was smaller than the number of boys. In the remaining two classes where the girls scored lower, the number of girls was greater than the number of boys. Apparently the sex group that was in the majority in the classroom was more motivated by the class spirit and endeavored to raise the class score even at the cost of honesty.*

Summary

An attempt was made to study children's behavior on an honesty test under personal and social motivation. Additional scores of

* See "Studies in Deceit," by Hartshorne and May, Chapter XXII, The Motivation of Deceit, where this finding is quoted as an illustration of the effect of motivation upon deceit.

honesty and cooperativeness were available. The results may be summarized as follows:

1. In general children were more deceptive where the score on the test was to count for a personal gain than when the score was to count as a gain for the group.

2. Similarly, the children worked at a much higher efficiency when working for self than when working for the group.

3. There were, however, marked individual differences in this respect. There were some individuals who cheated more in their group work than in self work. Similarly, there were a few children who worked at a higher speed when working for the group than when working for themselves.

4. There was a positive correlation between cooperation and deception in group work. Those who scored high on the cooperation test tended to cheat in favor of their group work more frequently than those who scored low on the cooperation test.

5. There was a low but negative correlation between cooperation and deception in work for self.

6. There was a marked positive correlation between dishonesty in self work and dishonesty in group work.

Apparently some children learn to be more cooperative even at the cost of honesty. Others learn to be honest even at the cost of some personal or group gain. When there is no conflict the correlation between cooperation and honesty is somewhat positive. Those who act honestly on an honesty test tend also to be cooperative on a cooperation test. But when honesty and cooperation are thrown into conflict, *i.e.*, when one may add to the gain of the group by means of dishonesty, the correlation becomes definitely negative. An individual's behavior in situations of this nature seems to be determined to a slight extent by a general tendency to act in accord or in discord with accepted standards, but to a large extent by the individual's specific training in the given situations.

A LABORATORY STUDY OF THE READING OF FAMILIAR FORMULAS

G. NEVIN REBERT

Hood College

The Problem

This investigation was made to determine how familiar formulas are read when they appear in verbal context. The investigation was concerned with the differences, if any, which exist between the eye-movements made during the reading of words in context and those made during the reading of familiar formulas in context. It aimed to discover whether proficiency, on the part of the reader, in respect to the subject-matter has any effect upon the character of the eye-movements made in reading formulas.

Two previous eye-movement studies deal with the reading of formulas. In connection with a comprehensive study of various types of silent reading Judd and Buswell[1] present records of the reading of a passage which contained several algebraic expressions. The records show a tendency on the part of the subjects to read the algebraic expressions in detail with a relatively greater number of fixations than was used to read the words. Tinker[2] made a laboratory investigation of the reading of algebraic and chemical formulas in verbal context and the reading of algebraic formulas in isolation. Tinker found that there is a tendency for the presence of the formulas to increase the number of fixations per line. He says that the mental processes used in reading formulas differ from those used in reading prose. He concludes that the nature of the eye-movements is the same whether the reader has a little, a moderate amount, or much training in the subject-matter. Concerning the reading of chemical formulas, Tinker says that a single fixation near or on it can mean only that the formula was looked at but not read.

The present investigation was concerned primarily with the reading of *familiar* formulas. Certain formulas occur so frequently in textual material that it seemed desirable to determine what modifications, if

[1] Judd, Charles Hubbard, and Guy Thomas Buswell: Silent Reading: A Study of the Various Types. *Supplementary Educational Monograph*, No. 23. Chicago: Department of Education, University of Chicago, 1922.

[2] Tinker, Miles A.: A Photographic Study of Eye Movements in Reading Formulae. *Genetic Psychology Monographs*, Vol. III, No. 2, 1928.

any, occur in the eye-movements of those who have occasion, frequently, to read them. Illustrations are $a^2 + 2ab + b^2$ in algebra, H_2SO_4 in chemistry, and E.M.F. in physics. This investigation was made to determine how such formulas are read in verbal context.

Procedure

The Subjects.—Sixty subjects participated in the investigation. They were selected on the basis of the degrees of efficiency they had attained in one or more of three sciences—chemistry, mathematics, and physics. Novice, advanced, and expert subjects were used.

Thirty-one subjects were used in the investigation of the reading of chemical formulas. The novice group consisted of six subjects without prior training in chemistry who had just completed one-half of the first course in inorganic chemistry in college. The advanced group was composed of sixteen subjects who had just completed one-half of a course in inorganic chemistry in college which had as prerequisite a course in chemistry in high school. The expert group was made up of nine subjects who had completed at least two years of graduate work in chemistry. Two were Doctors of Philosophy and two were Masters of Science in chemistry.

Thirty-one subjects used in an investigation of the reading of algebraic trinomials were divided into three groups according to their training in mathematics. The novice group consisted of five high school seniors who had done outstanding work in general high school mathematics. The advanced group was composed of fifteen subjects who were familiar with calculus. The expert group consisted of eleven advanced graduate students in mathematics or teachers of mathematics with two or more years of graduate training. Twelve of the mathematics experts read two calculus selections.

Seventeen subjects were used in the investigation of the reading of physics formulas. The novice group was comprised of five college freshmen who had received credit for a course in physics in high school. The advanced group consisted of four subjects who had completed a course in physics in college in addition to the high school course. The expert group was composed of eight subjects who had completed at least two years of graduate work in physics.

The Reading Selections.—Six of the reading selections used in the investigation are shown in Plates I to VII. In addition an American history selection was read by each subject as a check upon his general reading ability. This selection is shown in the report of a previous

investigation by the writer.[1] The selections were printed in lines four inches in length excepting the one shown in Plate IV which was printed in lines 3½ inches long. The spacing between the lines varied from 1/32 to 9/32 of an inch in width. Twelve point type was used.

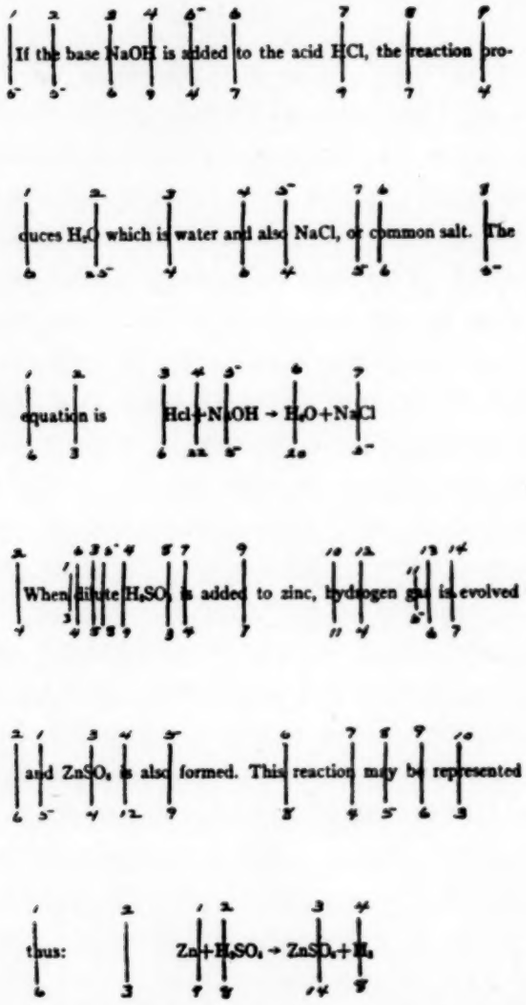

If the base NaOH is added to the acid HCl, the reaction pro-
duces H2O which is water and also NaCl, or common salt. The
equation is Hcl+NaOH → H2O+NaCl
When dilute H2SO4 is added to zinc, hydrogen gas is evolved
and ZnSO4 is also formed. This reaction may be represented
thus: Zn+H2SO4 → ZnSO4+H2

PLATE I.—Eye-movement record of a chemistry expert. The vertical lines indicate the point of fixation. The serial numbers above these lines indicate the order of the fixations; the numbers below the lines, the duration of fixations in units of twenty-fifths of a second.

In the chemistry selection (Plate I) the formula, HCl, was printed, Hcl. In the algebra selection (Plate III), $x^2 + 2xy + y$ was termed a perfect square. These errors were purposely inserted in the text of the two selections in order to determine the amount of detail with which these familiar formulas were read.

[1] A Laboratory Study of the Reading of Familiar Numerals. *Journal of Educational Psychology*, Vol. XXIII, No. 1, pp. 35:45.

Techniques of the Investigation.—The apparatus used in the investigation was the eye-movement camera in the laboratory of the School of Education at the University of Chicago. The one hundred three film records of the readings of the selections by the sixty subjects were interpreted by the same method that has been standardized in pre-

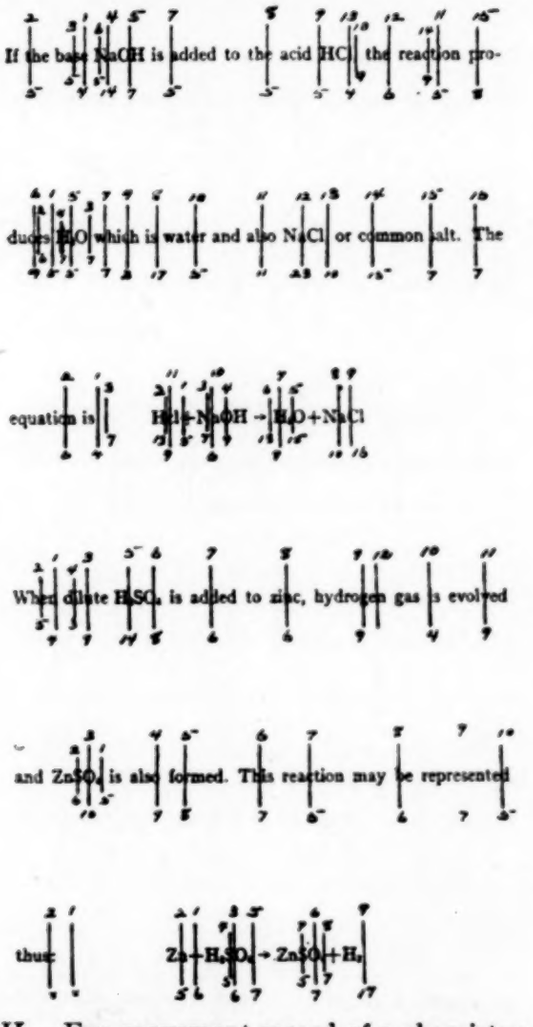

If the base NaOH is added to the acid HCl the reaction pro-
duces H2O which is water and also NaCl or common salt. The
equation is
HCl + NaOH → H2O + NaCl
When dilute H2SO4 is added to zinc, hydrogen gas is evolved
and ZnSO4 is also formed. This reaction may be represented
thus:
Zn + H2SO4 → ZnSO4 + H2

PLATE II.—Eye-movement record of a chemistry novice.

vious investigations. A description of the method of determining which fixations fell upon the formulas may be found in the report of a previous investigation by the writer.[1]

RESULTS

Proficiency in Subject-matter and Efficiency in Reading.—The measurements of eye-movement records of the novice, advanced, and expert groups in chemistry in the reading of the familiar historical

selection and the chemistry selection (Plate I) are shown in Table I. The records of the reading of the historical selection are used as standards with which to compare the measurements of the eye-movement records made in reading chemical formulas. This is appropriate because these measures show the ability of the three groups to read familiar material, other than material chemical in nature.

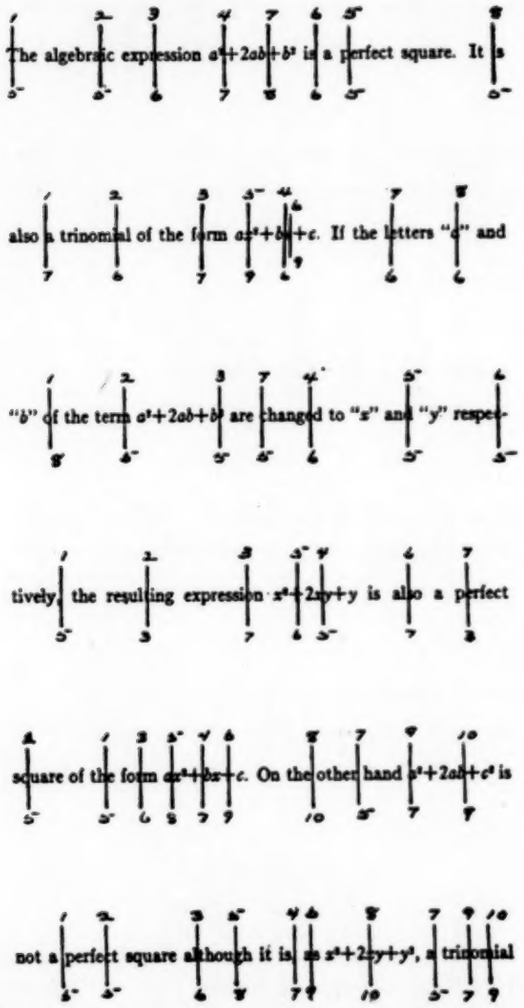

The algebraic expression $a^2+2ab+b^2$ is a perfect square. It is
also a trinomial of the form ax^2+bx+c. If the letters "a" and
"b" of the term $a^2+2ab+b^2$ are changed to "x" and "y" respec-
tively, the resulting expression $x^2+2xy+y$ is also a perfect
square of the form ax^2+bx+c. On the other hand $a^2+2ab+c^2$ is
not a perfect square although it is as $x^2+2xy+y^2$, a trinomial

PLATE III.—Eye-movement record of a mathematics expert.

It may be observed in Table I that in respect to the history selection there appears to be no significant variation between the three groups of records in the mean number of regressions per line. Taking these records as a whole, however, the expert group seems to read slightly more rapidly than the novice and advanced groups. On the other hand, the records of the reading of the chemistry selection show that the groups more highly trained in chemistry tended to read this selec-

TABLE I.—EYE-MOVEMENT RECORDS OF NOVICE, ADVANCED, AND EXPERT GROUPS IN READING A HISTORY AND A CHEMISTRY SELECTION

Group	Number of subjects	Mean					
		Duration[1] of fixations		Number of fixations per line		Number of regressions per line	
		History	Chemistry	History	Chemistry	History	Chemistry
Novice	6	6.8	7.6	8.4	13.8	1.6	3.1
Advanced	16	6.3	6.5	8.3	9.8	1.4	1.6
Expert	9	6.4	5.8	7.8	9.3	1.3	1.9

[1] Time in units of $\frac{1}{25}$ seconds.

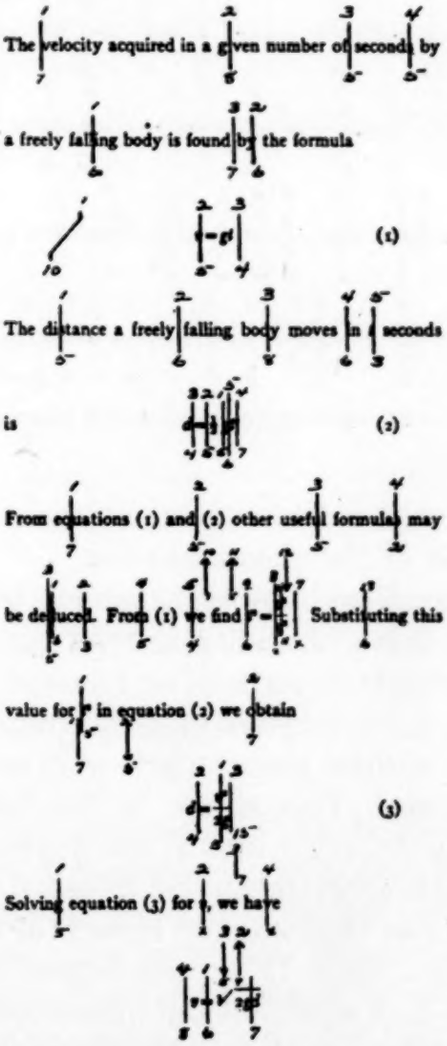

The velocity acquired in a given number of seconds by
a freely falling body is found by the formula
(1)
The distance a freely falling body moves in t seconds
is
(2)
From equations (1) and (2) other useful formulas may
be deduced. From (1) we find
Substituting this
value for v in equation (2) we obtain
(3)
Solving equation (3) for t, we have

PLATE IV.—Eye-movement record of a physics expert.

tion with markedly fewer fixations and fixations of markedly shorter duration than did the subjects less highly trained in chemistry. The advanced and expert groups made markedly fewer regressions per line than the novice group in reading the chemistry selection. However, the *advanced* group made relatively fewer regressions than the *expert* group.

The records of the reading of the algebra and physics selections support the findings concerning the reading of the chemistry selection, namely: Subjects more highly trained in content tend to read the selections with relatively fewer fixations, fixations of relatively shorter

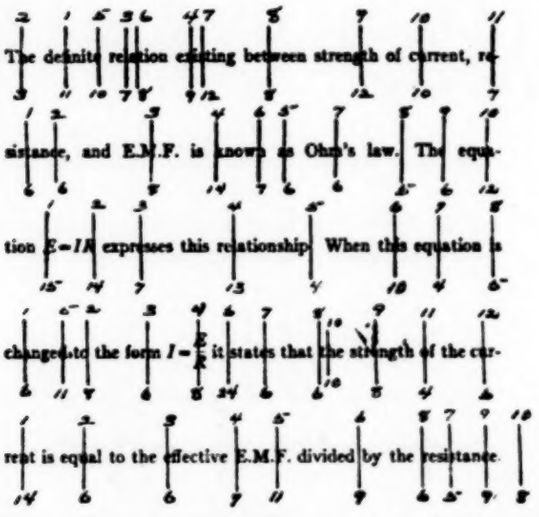

The definite relation existing between strength of current, re-
sistance, and E.M.F. is known as Ohm's law. The equa-
tion $E=IR$ expresses this relationship. When this equation is
changed to the form $I=\frac{E}{R}$ it states that the strength of the cur-
rent is equal to the effective E.M.F. divided by the resistance.

PLATE V.—Eye-movement record of a physics expert.

duration, and relatively fewer regressions than subjects less highly trained in the content of the selections read.

The number of fixations, however, cannot be used in individual cases as an accurate index of ability to read mathematical formulas. In the case of the algebraic trinomials, for instance, there were readings with one fixation made by subjects in each group—novice, advanced, and expert. In each of these groups there were readings that involved seven or more fixations. This is due to the fact that there was a tendency for the subjects, regardless of the degree of training, to read the algebraic trinomials analytically and in detail rather than as units. This same tendency was shown in the records of the reading of one of the physics selections (Plate IV). One expert subject consistently made one fixation on each equation and understood what he had read. The other expert subjects tended to read this selection analytically

and in detail. The record shown in Plate IV is typical of such analytical reading.

Comparison of the Measurements of Eye-movement Records Made in Reading Formulas and Their Verbal Contexts.—Table II shows the mean duration of the fixations made by sixty subjects on the verbal contexts and formulas of six selections during a total of one hundred three

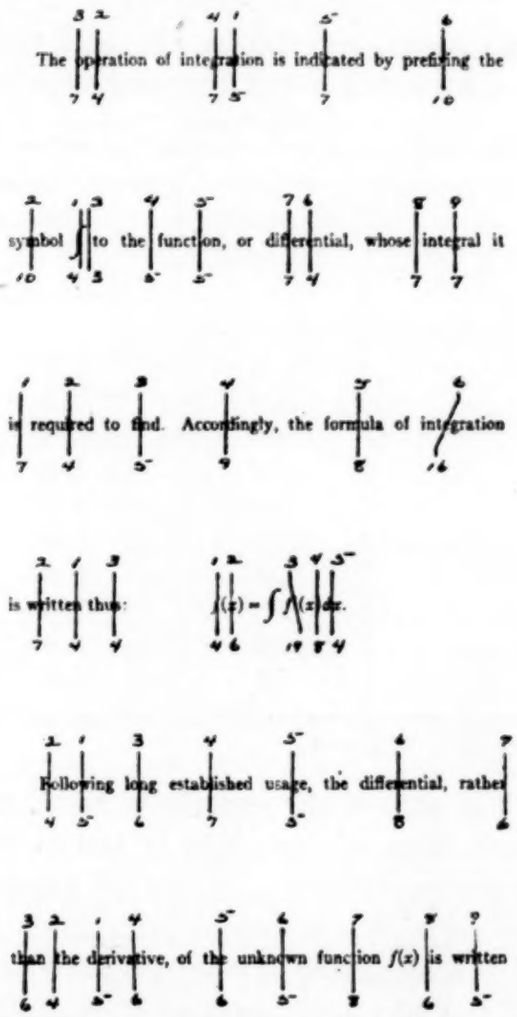

The operation of integration is indicated by prefixing the
symbol ∫ to the function, or differential, whose integral it
is required to find. Accordingly, the formula of integration
is written thus: ∫f(x) = ∫f(x)dx.
Following long established usage, the differential, rather
than the derivative, of the unknown function f(x) is written

PLATE VI.—Eye-movement record of a mathematics expert.

readings. It may be observed in Table II that in respect to each of the six selections read, fixations of longer duration tended to be made upon the formulas than upon their verbal contexts. In nine only, of the one hundred three individual records on which the measures in Table II are based, was the mean duration of fixations upon the verbal context greater than that upon the formulas.

Tendency to Read Formulas as Units.—Sixty-seven per cent of the readings by chemistry experts on the six formulas which appear in the verbal context of the chemistry selection were unitary. Forty-eight per cent of the readings of these formulas by the subjects of the advanced group were of the same character. None of the five novices in chemistry could read the formulas as units. (Plate II shows a relatively efficient novice reading record.)

TABLE II.—MEAN DURATION OF FIXATIONS MADE BY SIXTY SUBJECTS IN ONE HUNDRED THREE READINGS OF THE CONTEXTS AND FORMULAS OF SIX SELECTIONS

Selection	Number of subjects	Mean duration of fixations in units of 1/25 seconds	
		Verbal context	Formulas
Algebra	5[1]	7.2	8.1
	15[2]	7.0	8.3
	10[3]	6.3	7.1
Chemistry	5[1]	7.6	8.6
	16[2]	6.5	7.9
	9[3]	5.8	6.8
Calculus I	9	5.8	7.2
II	6	5.9	6.7
Physics I	12	6.6	11.0
II	16	6.0	7.0

[1] Novice. [2] Advanced. [3] Expert.

A tendency on the part of the experts in mathematics and physics to make unitary readings of formulas is indicated by the reading records. This tendency, however, is not so pronounced as that in the case of the reading of the chemical formulas. Eighteen per cent of the readings of the algebraic trinomials and thirty-nine per cent of the readings of the physics formulas were unitary in character.

The records of the reading of the chemistry selection by the nine chemistry experts indicate that it is possible to read the formulas in detail with one fixation. Seven of the nine experts made only one fixation upon HCl. Four of the experts recognized the error in Hcl. Seven of the nine experts did not fixate NaCl more than once. Six of the nine experts fixated H_2SO_4 once. Each of these subjects was able to enumerate the six formulas after he had read the selection. The fact must not be overlooked, however, that there was also shown

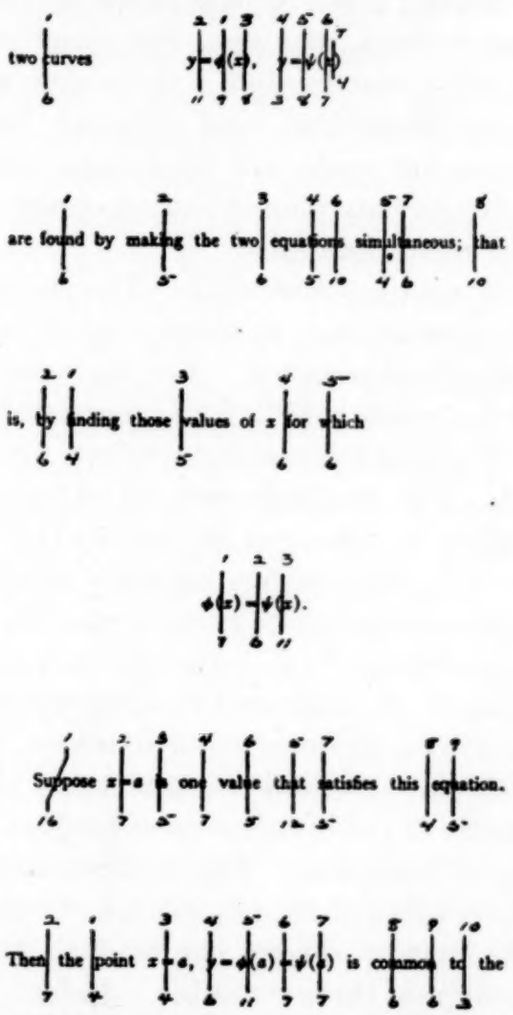
a tendency to fixate these formulas more than once. Each of the formulas was fixated more than once by one or another of the nine expert subjects. Five of the nine subjects fixated both NaOH and $ZnSO_4$ more than once. Since familiar words of from three to five letters in context are rarely fixated more than once by mature readers,

two curves
are found by making the two equations simultaneous; that
is, by finding those values of x for which
Suppose x = a is one value that satisfies this equation.
Then the point x = a, y = φ(a) = ψ(a) is common to the

PLATE VII.—Eye-movement record of a mathematics expert.

it seems evident that chemical formulas are more difficult to read than words.

The advanced and expert physics subjects tended to read E.M.F. as a unit with one fixation. Thirteen out of twenty readings were unitary in character. In four cases, as in the record shown in Plate V, a subject made one reading with one fixation and the other with two fixations. Seven of the twenty readings involved two or more fixations.

Six expert mathematicians read the calculus selection shown in Plate VII. The equation, $x = a$, relatively simple in form, was fixated once each in ten of twelve readings. The two equations, $y = \phi(x)$ and $y = \Psi(x)$, were read with on fixation in three cases and with two fixations in six cases. On the other hand, the equation $\phi(x) = \Psi(x)$ was fixated three or four times by four of the six expert subjects. It seems evident that only the simplest form of equation tends to be read with one fixation. Although $\phi(x)$ and $\Psi(x)$ are very familiar to mathematicians, and although the context in which they appear prepares the reader for their appearance, the equational relationship in which they are placed tends to cause them to be directly fixated—frequently more than once.

Contrasts in the Reading of Formulas.—The character of the context in which a formula appears may so change conditions that an attitude of analytical reading is demanded. For instance, the formula $f(x)$, as it appears in verbal context (Plate VI) merely expresses the general idea of function. There is no reason, therefore, for the mathematician to read it in detail. The readings were all unitary in character. On the other hand, when it appeared in the context of the formula of integration (Plate VI), the mathematician's attitude toward it was changed. It was necessary that he note the specific functional relation expressed. Each character had a specific meaning. Consequently he assumed an attitude of analytical reading which was evidenced by increases in the number and duration of fixations.

The Effect of the Interpretative Context upon the Reading of Formulas.—The character of the interpretative context plays an important part in the reading of formulas. This is illustrated by the eye-movement records of the reading of the physics selections (Plates IV and V).

In Plate IV the context directs the reader's attention specifically to the interrelationships of three formulas. In fact, these interrelationships are the central concern of the selection. The context is merely a contributory element to their understanding. Consequently, the attention of the reader is directed to the details of the formulas rather than to the thought of their context. A high degree of analytical reading results.

In Plate V, on the other hand, the verbal context is such that the formulas are merely contributory to the thought, or, at the most, a restatement of it. In consequence, the reading of the formulas in this selection is not so analytical in character as that of the formulas of the selection shown in Plate IV.

Conclusions

1. Increase in ability to read formulas tends to be associated with decreases in the number and duration of fixations and in the number of regressive fixations made in reading them.

2. Familiar formulas tend to be read with a relatively greater duration of fixations than their verbal contexts.

3. Familiar chemical formulas tend to be read as units in the same manner as words are read although analytical readings occur. Novices in chemistry cannot read chemical formulas as units.

4. Formulas of a mathematical type tend to be read by the expert subjects analytically and in detail although unitary readings occur as in the case of chemical formulas. The analytical reading is probably accounted for by the fact that the primary function of the mathematical formulas, as used in the present investigation, was to express relationships rather than single concepts.

5. The character of the reading of formulas tends to be affected by the character of the context in which the formulas appear. When the formulas are merely contributory to an understanding of their contexts, they tend to be read as words are read. When the contexts are merely explanatory and direct attention to relationships between component characters of the formulas, the formulas tend to be read analytically and in detail.

TIMING AS A PHASE OF SKILL

COLEMAN R. GRIFFITH

University of Illinois

Most of the studies on learning and skill have taken it for granted that a skill is a pattern of bodily movement notable either for its absolute perfection or for the maximal speed at which it may be executed. Many types of juggling, skill in throwing at a target, many types of piano playing, and much machine work, are examples of skills wherein emphasis is laid upon a high degree of perfection in the form or pattern of movement. This phase of skill may be described by a word which comes from the athletic field, *viz.*, stance. Athletic stances are of two kinds. One may speak (1) of the initial or starting stance which is a static movement system or a posture (Washburn), and (2) of the action stance, that is, of the way in which any given act should be performed, and this is a phasic movement system (Washburn).[1] A golfer, for example, takes an initial stance which favors a precise swing of his club; and then he carries out an action stance which is pictured in the way he raises his driver, in his swing, in his follow-through, and so on. It is said by the golfer that there is a "way-in-which" the act of driving should be done and that this "way-in-which" may be called an action stance. The action stance of the golfer lays emphasis upon the degree of perfection of a series of bodily movements. Skill is judged by the criterion of accuracy.

There are many skills, however, which lay emphasis upon the sheer rate at which an act is to be executed. The card sorter attempts to distribute his cards just as fast as he can. The typist, the telegraph dispatcher, and the subject who is in the midst of a serial reaction, are all attempting to move just as fast as possible and the criteria of excellence or of degree of learning in studies of this kind are based solely on speed. Speed has become embodied in the literature of learning as "practical limit" or even as "physiological limit." It is observed that, as a man goes through a series of coordinated movements with greater and greater speed, he approaches a limit, or a rate beyond which he does not find it easy to go. This limit is commonly believed to be unsurpassable because it is found that most men do not, as a matter of fact, pass it. We do not as yet know, however, why

[1] Washburn, M. F.: "Movement and Mental Imagery." 1923.

men cannot excel alleged practical and physiological limits. It is lightly believed that nervous systems and muscles cannot work any faster and in the case of a hundred yard dash done at the rate of ten yards per second (or slightly better), this is probably true, as Hill[1] has shown; but it has not been shown that this can be true of other more complicated series of reactions, and especially of those reactions commonly known as mental. In the case of mental reactions, abortive attempts have been made to correlate speed of movement or even speed of nerve conduction with general intelligence.[2] In these studies and in studies which seek to determine the "practical" or even the "physiological" limits of speed in work it is assumed that any performance, when done under the instruction to work or to move as rapidly as possible, actually yields a rate which is maximal for that particular task.

In addition to skills that are judged by the two criteria of perfection and of maximal "rate-at-which" or speed, there are skills which are judged by both of these criteria. In some types of piano playing, for example, the nature of the composition is such as to demand both perfection and speed. Many of the compositions of Liszt are of this type. One must learn not only the pattern of movements symbolized in the score but one must acquire the ability to render this pattern of movements at great speed. In a sense, typewriting and dispatching are of the same type. One must touch the keys accurately but speedily as well. The rat must run the maze without going into blind alleys and without losing any time on the way. The fastest and most accurate rat is said to be the most skillful or the most learned. So far has theory of learning and laboratory practice gone in the direction of accuracy and speed that these two factors have entered into the preliminary instructions of at least ninety per cent of the so-called mental or alertness tests; and, as we have seen, many alleged intelligence tests make them, and especially the speed factor, a foundation for judgments concerning intelligence and for prophecies concerning character.

There is, however, another phase of skill or of learning where the "rate-at-which" or the speed of a movement acquires new significance. In many athletic skills and, as we shall see, in many skills outside of

[1] Hill, A. V.: "Muscular Movement in Man." 1927.

[2] McFarland, R. A.: The Rôle of Speed in Mental Ability. *Psychological Bulletin*, Vol. XXV, 1928, pp. 595–612.

athletics, the "rate-at-which," the timing, the pace of a skill must be neither minimal nor maximal but optimal. The rate that is chosen is of dominant importance in any criteria of excellence under which the skills may be placed. Consider, for example, the runner. He takes an initial stance which favors a quick start from his marks; the initial stance is transformed at the sound of the gun, into a running stance where such items as bodily position, length of stride, height, or pumping action of legs, swing of the arms, and the like, are prominent features. But another phase of the skill is even more significant, *viz.*, the rate at which a movement is to be executed. In a hundred yard dash this "rate-at-which" is almost a maximal rate. The track man runs with the best form and at the greatest speed it is possible for him to muster. But in all runs over a quarter mile (and sometimes in the two hundred twenty yard dash) the "rate-at-which" is not a maximal rate. It is an optimal rate, a pace which, for any given individual, is just as important as any other feature of his skill. A mile runner is said to have perfect pace or timing when he runs just fast enough to cover the whole distance within, let us say, four feet thirty inches, without wasting his energy in too fast a start or in useless sprints on the way. If, on the contrary, he runs too slowly, he will lose out on the final sprint. That this factor of timing must be brought to a high level of accuracy is shown by the fact that, on a track with six laps to the mile, a mistake of more than two or three seconds in running any lap, will put an almost impossible handicap on any man who is running against close competition. That this factor can be acquired and highly developed is shown by the experience of Joie Ray who could run a mile in almost any race within a second of the time imposed upon him by himself or by the speed of his competitors. Any track coach who has taken the pains to develop this aspect of running skill in a group of men will testify to the same fact. That is, coaching a runner is partly a problem of finding a rate of movement, a pace, a running stance, which is the optimal stance for any given runner, and the standard of excellence in this type of skill is based neither upon mere physical form nor upon a maximal rate, but upon this optimal rate.

There are plenty of other illustrations of timing or pace as a factor in skill. Let us take another case in athletics, *viz.*, the skill of the punter. There is a way to hold the ball, a manner of stepping forward after the ball is passed back from center, and a way of following through

with the foot while the head is still down that will insure a long, high punt. Many men practice this form of punting, this action stance, until they are perfect in doing it; but this does not at all mean that they will punt well in a game, as our observations and experiments abundantly show. During a game a punter has, on the average, 1.2 seconds available for getting his punt away. That is, he can hold the ball in his possession for that length of time. If he takes 1.6 seconds, approximately a quarter of his punts will be blocked (assuming that the forwards of the two teams are about equal in charging strength and in skill and that the punt is fairly blocked). If the punter takes as long as two seconds, he has practically no chance at all of getting his punt into the air. Now it appears that most men, when they practice punting during the week, lay all of their emphasis upon the form of their punt and acquire a rate of movement or a timing, which, from individual to individual, runs anywhere from two to four seconds. That is, during the week before a game most punters are busily engaged in acquiring skill in the form of punting, and they are also acquiring skill in a "rate-at-which" they should get the ball away. This "rate-at-which" varies from individual to individual; but within a week's time, unless steps are taken to correct the error, the timing habit will be well learned. The disastrous effects of this fortuitously acquired pace appear only during the game when a fast charging line allows the man but 1.2 seconds, or less, to get his punt away. In order to meet game conditions the punter has to break a habit. His punt, as we say, is hurried; it goes badly. It looks unskillful or even awkward. As a matter of fact, it is just as unskillful as it would have been if, at the last moment, the man had been obliged to make a fundamental change in the form of his movements rather than in the timing. It is as though a pianist should suddenly attempt to play "Rock of Ages" with the timing and rhythm of "Rainbow Round My Shoulder."

Nowhere does the distinction between form or pattern of movement and rate or timing appear more clearly than in boxing and nowhere do we find a better illustration of the effect of plateaus, practical limits, and retrogression than here. The form or the pattern of the one-two punch, for example, calls for a feint with the left arm, a quick shift of the feet, and a rebalancing of the weight of the body so that the right arm can be brought effectively against the jaw or body of the opponent. Skill in the form or pattern of this particular punch can be easily acquired by practicing upon a punching bag or punching

dummy. One can also practice it in what is called shadow boxing. But the greatest skill acquired by these means does not augur an effective use of the one-two punch in actual combat for there a new factor arises, *viz.*, the timing of the whole pattern. This timing is not merely the result of a judgment based upon the motions or the position of the opponent. It is based upon the fact that the opponent will react to the feint of the jaw in such a way as to leave him wide open for the second and punishing phase of the whole pattern. A man whose timing has gone wrong is just as inadequate in the ring as is the man who has lost his judgment of distance.

We gather from these observations and illustrations that timing or pace is a distinct factor in certain skills and that it is not to be confused with perfection in form or with maximal speed.

It is the purpose of the present study, therefore, to suggest that some so called maximal speeds, whether at the practical limit or not, may not be maximal at all, that instead they represent nothing more than a timing or pace habit that has been set as an artificial barrier against further progress. It is also our purpose to point out that many skills cannot be judged at all in terms of speed or even in terms of accuracy because the pace or the timing of the skill stands as the significant feature of it.

We have sought to put this phase of skill under experimental control in the following way. The obvious mode of procedure would have called for a method of training runners to practice a pace or rate of movement set before them in a mechanical way and then to have tested the rate of acquisition and rate of relearning of this pace by methods commonly used in similar researches. We have, as a matter of fact, adopted this method in a series of studies on swimmers, and it has been found possible to improve greatly their judgments of pace, but it has seemed that the problem ought to be opened up as a general psychological problem, related, if possible, to the studies in the field of rhythm and, as we shall try to show at the end of the paper, brought to bear in a practical way upon certain traditions in the field of learning.

Use has been made, therefore, of the slot maze for the study. The Carr maze had been modified in the following way.[1] Along the bottoms of the runways in a maze 60 × 120 centimeters, there was

[1] Carr, N. A.: *Journal Experimental Psychology.*

placed, at intervals of ten centimeters, a series of brass contact points over which a stylus could slide with no perceptible halting or jolting. These contact points find a common ground in a large grass plate upon which the whole maze rests. The stylus furnishes the other pole of a circuit which is connected through a Renshaw polygraph to the brass plate upon which the maze rests. The contact points in the alleys of the maze are so arranged that the stylus can never traverse more than ten centimeters without making a contact. This is true not only of the correct pathway but of the blind alleys as well. As the learner goes through the maze a record is kept, not only of the pathway pursued but of the rate at which any path has been traversed. These records are balanced against a record which is brought to the recorder from a second's pendulum and from specially constructed contact devices.

The problem has taken several different forms. One group of subjects, for example, was instructed to learn the maze at a rate of his own choosing. The only restriction was that the subjects of this group should make a subjective determination of a rate which seemed most natural to them. In their own language they were asked to choose a rate which they "instinctively preferred." They were told to attempt to perfect this rate as they traversed the maze. In the meantime, they were also perfecting the form of the habit, that is, the pattern of movements necessary to get through the maze. A second group of individuals was asked to learn the form of the maze under the same conditions that obtain in any experiment upon maze learning. That is, there was no restriction or comment made about pace save that they were told to move as quickly but as accurately as they could. When once the maze was learned, they were asked to impose upon the pattern and timing habits already acquired a series of timing habits where the rate was set for them by a second's pendulum which activated a buzzer. A third group of subjects was shown the correct pathway so that they could traverse it with a minimal amount of practice. Usually this process of learning took but three or four trials and this was not deemed enough to allow them to acquire any timing habit with respect to the pattern of the maze. When the pattern of movement necessary to pass through the maze was fully learned they were asked to go through the series, some to cover each unit distance (ten centimeters) in two seconds, some to cover it in one second, and some to cover it in a half second. There were thus three different

paces or timing habits to be acquired by these three different groups of learners.

The degree of attained skill was measured (1) by the number of trials it took to learn the form or the pattern of the maze, that is, the correct pathway, (2) by the number of trials necessary to learn the rate at which the maze should be run, and (3) by the degree to which the records, gained after a four weeks interval, paralleled those gained at the end of the initial learning period.

When subjected to the modes of statistical treatment adequate for this kind of an experiment, one fact stood out clearly enough, *viz.*, that timing or pace appears as a distinct factor in or phase of skill. This factor, at least in maze experiments, appears to be learned more slowly than form or pattern of movement and is forgotten more quickly. This is true under all of the conditions laid down save where the learner chose his own rate of movement—that is, that rate which seemed most natural to him. Since all of the subjects were athletes and all had had more or less experience in learning rates of movement in their several sports and since the chances are that each subject chose some rate or pace with which he was already familiar, this phase of the program must be pursued with further variations of conditions and with subjects who do not, perchance, have an acquired preference for some one pace. It was found, furthermore, that fast paces were much more difficult to acquire than slower paces. Finally, it was discovered that prodigious individual differences, exceeding even those that were found where we were concerned only with pattern of movement, characterize timing habits. At one extreme stood one subject who could go through the maze pattern as though he heard the ticking of a watch and at the other extreme stood individuals who appeared to have practically no sense of timing. Had we been selecting our data as animal experimenters often do, who refuse rats which cannot learn a maze, we should have thrown out these individuals. At any rate, the individuals' differences are there. It has not yet determined how much of this difference is due to training and how much to the anciently alleged intuitive sense of timing.

There is at least one serious fault in the study. We have laid down no plan for equating the difficulty involved in learning the form or pattern of a maze against the difficulty involved in learning a pace or a rate of movement. That is, we do not know how difficult the pattern of a maze ought to be in order that it shall stand upon the same

level as pace or timing. There is only the general suggestion that it is exceedingly difficult to learn a pace. Track men work for months upon the problem and some of our subjects quite failed to approach the levels of skill which we thought they might or ought to approach. An initial series of experiments failed to give us all that we had hoped for because the maze was so simple that there was no chance to compare the curves of learning for form or pattern with the curves for pace or timing. A second maze, like the first, save that it was many times more difficult, was constructed; and our belief is that this provides almost as difficult a problem as is afforded by timing.

The criteria for equal difficulty are as follows. It may be assumed that if two fundamentally different types of learning yield learning curves which closely resemble one another, the problems must be approximately equal. This resemblance should be so great that a single equation might describe both curves. When a maze has been found which yields on the average a curve of learning for form or pattern of approximately the same character as it yields for pace or timing, it may be assumed that a way has been found to balance skill in pattern of movement against skill in pace or timing. In the new maze it looks as though we had approached this equality in problem. We can approach it because pace or timing stands at the same level of difficulty regardless of distance traveled, number of blind alleys, or what not. That is, having learned to move a unit distance during a given interval in a simple maze will enable a man, without further practice, to move a similar unit of distance in a given interval in a complicated maze; but having learned a simple maze does not contribute materially to learning the form or pattern of a larger maze.

These experiments ought to throw an interesting light upon one or two common beliefs about learning, especially when we take them in conjunction with a group of experiments on piano playing. Before the arithmetical process of smoothing has blotted out many of its distinguishing features, a great many learning curves show what have been called "plateaus." In some experiments on learning to play the piano there was reason to believe that the plateaus in learning were related in no uncertain way to the acquisition of fortuitous habits of timing or pace. That is, the attention of the beginners in piano was focused upon the task of converting the symbols of the score into a pattern of finger and arm movements. The learner inadvertently fell into a timing or pace habit which was stabilized along with the

pattern that was being acquired. Then, as other patterns were attempted, this previously acquired timing habit stood in the way of a new skill that was being learned where the criteria of speed was the sign of progress. There appeared, therefore, a plateau in the learning curve of this subject. In other words, the question is raised here again as to methods of instruction in piano. Many methods lay emphasis upon form at the expense of timing. When this is done, habits of pace or timing may interfere with further progress in other pieces of music, whenever emphasis is laid upon speed.

Similar considerations obtain with respect to what is commonly known as the practical limit of learning and modes of distinguishing the practical from the physiological limit. It is said that the practical limit is the point in the acquisition of a skill beyond which it is not easy to go because the return in skill per unit of time spent in practice is so small as to take away all incentive for further practice. In other words, the practical limit of learning is a plateau that stands moderately close to the upper limit of human talent. But there is another way in which the practical limit can be considered. A close examination of some of the learning curves found in several different types of athletic skill show that the practical limit is often set by the fact that the subject has fallen into a pace of working and not by the fact that he had acquired all of the skill that it is possible for him to acquire. In the simple observations mentioned above, punters on the football field definitely fell into pace habits which interfered with their skill whenever they were forced to set for themselves a faster pace. Analogous studies of differences in rate of movement as one shoots at a basket in basketball show that men usually, rather than rarely, fall into pace habits and that when game conditions call for a faster tempo of muscular coordination, their skill decreases.

In typewriting, for example, a person may meet his practical limit at forty words per minute. When he attempts to write at the rate of fifty words per minute errors multiply, not because the pattern of movement is suddenly changed but because the pace is changed. Errors multiply for the same reason that they would multiply if the subject changed from a standard keyboard to a newly arranged keyboard. One's problem in acquiring greater skill is not to learn the old keyboard better (assuming that one has learned the touch system) but to acquire a new pace. This task is not that of learning an old

pace better, or, as the saying goes, of making an old pace more automatic as we make old patterns more automatic; but it is a question of forgetting an old pace habit and acquiring a new pace habit. Early in the learning period this new problem introduces a plateau in the learning curve. Late in the learning period it occasions the practical limit and the removal of further incentive to learning.

ABSTRACT INTELLIGENCE AND ART APPRECIATION

HERBERT A. CARROLL AND ALVIN C. EURICH

University of Minnesota

Aesthetic appreciation is generally thought of as a psychological trait wholly removed from abstract intelligence. It is regarded as an emotional rather than an intellectual experience. If, however, the psychological principle that desirable characteristics tend to be positively correlated is accepted, the posit n that the two are completely separated is untenable. The questic en becomes: How high is the correlation? What effect does int ce of the sort measured by mental tests have upon appreciati

The following discussion will be limited to a consideration of the *appreciation* of art. No atte pt will be made to find the relationship between intelligence and *creative ability* in art for creative ability is obviously not identical with the ability to appreciate or to judge. It is more comprehensive; although probably including the critical faculty, it goes far beyond.

Measuring Instruments

Only recently have psychologists attempted to construct tools with which to measure appreciation. In the field of art, Norman Meier, at the University of Iowa, and Margaret McAdory, at Teachers College, Columbia, have succeeded in reducing this intangible variable to quantitative terms. The tests they have constructed are not wholly satisfactory, but do make it possible to speak objectively, and with a fair degree of reliability, of amounts and degrees of critical ability. In this study the Meier-Seashore Art Judgment Test and the McAdory Art Test were used to measure the amount of ability possessed by certain groups to appreciate art in order that it might be learned to what extent such ability was related to abstract intelligence.

According to Meier, the Meier-Seashore Test has a reliability coefficient of .71. This is not high, and cannot be explained away with the ease that Meier assumes. He writes in the Manual, "With tests based upon concrete learning accomplishment a higher reliability is expected than one testing complex mental functions. With the latter kind, a coefficient of reliability of .80 is regarded as about as high as can reasonably be expected, because of the uncertainity of knowing exactly what factors operate in the person's total reaction.

With a test of capacity the opportunity of chance factors to control the final result are increased, hence a somewhat greater allowance must be made for them."

This is pure rationalization. If a test of appreciation, elusive though that variable is, cannot be made as reliable as tests of other variables, then it is not a wholly satisfactory test, and should not be considered as such. With such low reliability, any statements concerning individuals would be most hazardous. This, however, does not hold true with respect to groups. Hence, it is felt that it may be safely used in a study of this kind.

To check Meier's reliability coefficient of .71, the present writers applied the Sherman-Brown formula to the correlation between the odd and even scores on the papers of one hundred thirty-five college students selected at random (see Table I). The estimated reliability coefficient was .78. A Pearson *r* was also found between the scores of forty-nine college students who took the test in 1929 and again one year later, in 1930. This reliability coefficient was but .61 ± .06. The second correlation may have been affected by the fact that some of the students had received considerable training in art during the year.

TABLE I.—RELIABILITY COEFFICIENTS FOR THE MEIER-SEASHORE AND MCADORY ART JUDGMENT TESTS

Measures	Meier-Seashore		McAdory	
	N		*N*	
Published *r*	69	.71	...	.93
Spearman-Brown formula	135	.78	100	.82
Repetition of test after one year for				
Art majors	49	.61 ± .06		
Non-art students	103	.69 ± .03		

Miss McAdory was much more successful than Meier in getting a high reliability. She publishes a coefficient of .93. This high degree of consistency is due in part to the length of the test since the time required for taking it is approximately three times that needed by the Meier Seashore Art Judgment Test. The test was validated on the basis of the consensus of the judgments of one hundred competent critics. Unanimity of opinion would have been infinitely more desirable, but probably quite impossible to attain. It appears to the writers that the disagreements of the critics might

have been cared for at least in part through a weighted method of scoring.

In a check of Miss McAdory's reliability coefficient, one hundred papers of college students yielded a Spearman-Brown coefficient of .82.

To discover whether or not the Meier-Seashore and the McAdory tests were measuring the same trait, a correlation between the two was calculated. The result was most surprising: .27 ± .06 (see Table II). This is a barely reliable correlation, which indicates that not only do the tests fail to measure the same kind of art ability, but also that the two strands which they do measure show practically no tendency to vary together.

TABLE II.—THE RELATION BETWEEN THE MCADORY AND THE MEIER-SEASHORE ART JUDGMENT TESTS

Tests	*N*	*r*	Corrected for attenuation
Meier-Seashore and McAdory.....	111	.27 ± .06	.34

INTELLIGENCE AND ART JUDGMENT ON THE COLLEGE LEVEL

In his Examiner's Manual Meier published a correlation of −.01±.09 between art judgment as measured by his test, and general intelligence, as measured by the Thorndike Mental Examination. The subjects were fifty-three college undergraduates.

Desiring to check Meier's findings the present writers correlated the intelligence ratings and the art rating of a number of students in the College of Education at the University of Minnesota. Intelligence was determined by means of the Miller Analogies Test,[1] which is given to all students entering the College of Education. In rating the art judgment ability of the students, both the Meier-Seashore and the McAdory tests were used. In each instance a slight positive correlation was found, although it was not a reliable one in the case of the McAdory test. (See Table III.)

ACHIEVEMENT AND ART JUDGMENT AT THE COLLEGE LEVEL

Another approach to the problem lies in discovering the relationships which exists between critical ability in art and achievement in college courses. The results of this type of analysis appear in Table

[1] Not commercially available.

IV. A comparison of the correlations between achievement and McAdory scores on the one hand and achievement and Meier-Seashore scores on the other is rather striking, the coefficients being identical in one instance and only slightly different in the other two. It is rather

TABLE III.—RELATIONSHIP BETWEEN ART ABILITY AND INTELLIGENCE ON THE COLLEGE LEVEL

Intelligence test	Art test	*N*	*r*
Miller analogy	Meier-Seashore	674	.26 ± .02
Miller analogy	McAdory	203	.10 ± .05

TABLE IV.—RELATIONSHIP BETWEEN ART ABILITY AND ACHIEVEMENT ON THE COLLEGE LEVEL

Achievement in	*N*	Meier-Seashore	*N*	McAdory
Art subjects	88	.26 ± .07	113	.24 ± .06
Non-art subjects	87	.15 ± .07	109	.15 ± .06
All subjects	88	.23 ± .07	113	.17 ± .06

surprising that the correlation between achievement by art majors in art subjects and the ability of the same groups to render art judgments is no higher than .26. This indicates, among other things, that success in routine art courses does not necessarily prove that a student possesses any considerable ability in appreciating art. Formal courses in art, like formal courses in literature, tend to stress information, even though information is but a single strand in appreciation and probably far from being the most important one.

The correlations do indicate, however, that there is a greater relationship between success in art subjects and art test rating than there is between success in non-art subjects and art test rating.

ART JUDGMENT ABILITY OF BRIGHT AND DULL CHILDREN

Even though the correlation between abstract intelligence and the ability to appreciate art is low it is possible for art judgment to be appreciably affected by extremes of intelligence. Obviously, the idiot does not possess a normal sense of aesthetic values.

To find out how groups at the two ends of the normal curve of intelligence compared in art judgment ability, forty-three backward children from Minneapolis opportunity classes and forty-three intellec-

TABLE V.—DIFFERENCE IN ART ABILITY BETWEEN BORDERLINE AND INTELLECTUALLY GIFTED CHILDREN

	N			CA				IQ				McAdory art scores			
	Boys	Girls	Total	Range	*M*	*PEM*	σ	Range	*M*	*PEM*	σ	Range	*M*	*PEM*	σ
Bright	19	24	43	11–6 to 15–0	13.37	± .08	0.74	119–154	132.24	± .87	8.44	104–194	158.49	± 2.20	21.44
Dull	19	24	43	13–10 to 16–9	15.26	± .09	0.92	61–80	71.05	± .57	5.50	69–178	101.28	± 2.59	25.15
Difference[1]	..	..	..	.....	+1.9	± .12	.....	.....	−61.19	± 1.04	.....	.....	−57.21	± 3.40	

[1] A minus sign indicates that the dull rate less than the bright; a plus sign, more.

tually gifted children from one of the best junior high schools were given the McAdory Art Test. The sex factor was equated in the composition of the groups, as it had previously been found that the girls earned a much higher art rating than the boys. As will be seen by Table V, the bright group was on the average approximately two years younger than the dull. If they had been exactly equated on the basis of chronological age, the differences would probably have been accentuated.

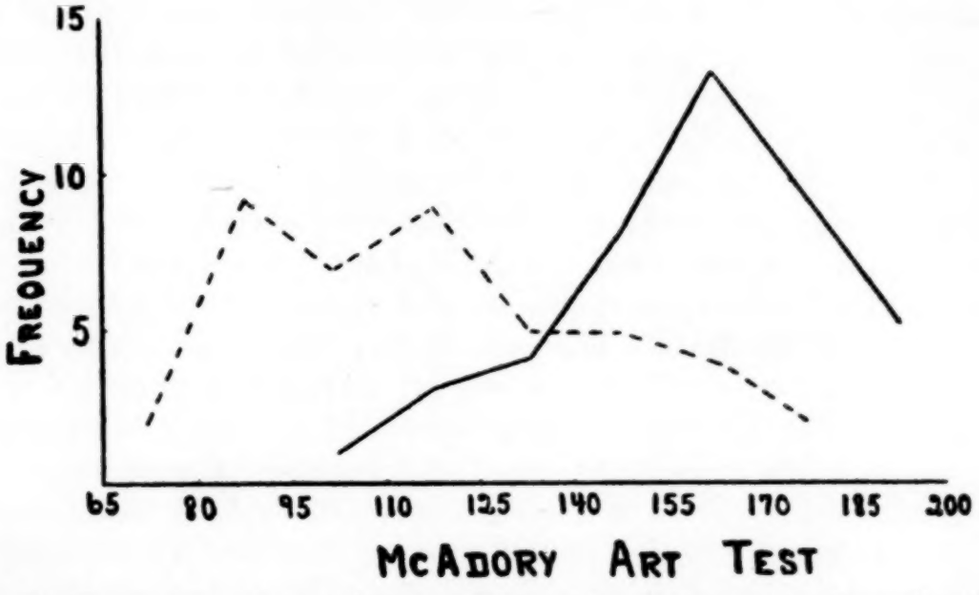


FIG. 1.—Comparison of art judgment ability of borderline and intellectually gifted children.

The dull group, with a mean IQ of 71.05 falls very near the line of feeble-mindedness, while the bright group with a mean IQ of 132.24 falls well above the gifted level. The difference between the intelligence of the two groups is, then, very marked, *i.e.*, 61.19 points. The reliability of the difference is high, giving a critical ratio of 59.

An outstanding fact about the comparative ability of these two groups on the McAdory Art Test is the considerable amount of overlapping (see Fig. 1). Although there is a decided difference between the means, *i.e.*, 57.21 points, still sixty-three per cent of the borderline children exceed the lowest score made by the gifted. The percentage exceeding the mean of the intellectually gifted group, however, is but nine, and the highest dull child is sixteen points lower than the highest bright child. No borderline child earns a score above plus sigma of the distribution of the gifted.

While nine per cent of the dull children were found to exceed the mean art score of the bright, no bright child falls below the mean and but twelve per cent below plus sigma of the borderline group. It would appear from this that it is somewhat more likely that a very dull child will possess considerable ability in art appreciation than that a very bright child will be seriously deficient in that ability.

Conclusions

The writers make no sweeping generalizations as a result of this investigation. Reducing appreciation to quantitative terms is a hazardous undertaking at best, and any findings presented at the present time must necessarily be tentative, since they rest upon measuring instruments which, as was pointed out at the beginning of this article, are not wholly adequate. However, the indications, briefly summarized, growing out of the data just presented are: (1) The correlation between the Meier-Seashore and the McAdory tests is positive but very low; (2) The McAdory Test appears to be the more reliable instrument; (3) Abstract intelligence has little if any relationship with critical ability in art on the college level; (4) Abstract intelligence at the extremes seems to effect art judgment ability. When it appears in excess, as in the case of the gifted, it precludes the possibility of an unusually low score; when there is a deficit, as in the case of the near feeble-minded, it does not preclude the possibility of a comparatively high score; (5) Gifted children are superior to borderline children in art judgment ability, but not so superior as they are in abstract intelligence.

ORDER AND FREQUENCY OF OBSERVATIONS RECORDED BY A GROUP

HENRY BOWERS

Ottawa Normal School

Although a considerable amount of both qualitative and quantitative work has been done on "observation" little attention has been paid to the mean order and mean frequency of objects observed by a group.

In this paper evidence will be presented showing that a tendency exists for groups of subjects to record the items observed in a picture (*a*) in the same order, (*b*) with the same relative frequencies (*c*) so that the rank and frequency vary together.

Procedure

The subjects used were High School pupils in every year of the course. Coloured pictures, "A Mountain River," "Wheat Harvesting in the Great North-West," and one in flat-tones of a Labrador Fiord were the test material.

In obtaining a "report" for each picture this procedure was followed. Every pupil was given a copy and a sheet of foolscap. The class was then addressed thus—"We are interested in how well you can observe things. To find this out you have been given a picture turned face downwards. Two minutes will be allowed for looking at this picture and then you will write down everything about it that you can remember. For each correct statement one mark will be given. Thus, if you say, 'a blue bird with a white feather in its wing'[1] and if each statement is true, a mark will be given for 'blue,' one for 'bird,' one for 'white feather' and one for 'in its wing.' Do your very best to make as many marks as possible. Turn the picture over. Go!"

(*At the end of two minutes*) "Turn the picture over. Write down everything you can remember. It is unnecessary to make sentences."

Each subject was given as much time as he required.

Whenever a "description" was desired, the "reports" were collected and the subjects requested to write down everything that could be seen in the picture. The reagents were permitted to use the picture while writing the "description." Again, no time limit was

[1] This was written on the blackboard.

set. Neither "report" nor "description" was supplemented by questions.

Notwithstanding the above example of scoring, the items in which the writer was interested in connection with this paper were purely substantive. For every "report" and "description" each *object* named was recorded and its rank attached. It was necessary to give conventional names to each object noted and these names only were used in the record. Thus, in the writer's opinion, there is a large stone in the middle of the mountain river. Although this was designated in the different reports as "rock," "boulder," "boat," "shipwreck," "log," and "stump," these were all translated as "stone" and credit given. A difficulty arose in connection with the repetition of certain objects. For instance, if there were two men, or three binders in a picture it would be impossible in every case to assign an order to each. In such cases where doubt might arise as to which of several items falling in the same category was referred to, credit was given, and order assigned, to the first one named. Others were neglected. These are serious difficulties incidental to the material employed but it is not considered that they are sufficiently grave to impair the validity of the results. On February 6 "reports" for the picture "Mountain River" were secured for eighteen subjects using an observation period of two minutes. On May 1 these eighteen were given the same picture for five minutes and again asked for a "report." In both cases attributives as well as substantives were scored. The Pearson r coefficient between the two sets of scores was $.86 \pm .05$. Whatever other inference may be drawn from this figure a fair measure of consistency in scoring is indicated.

The following were determined for each object[1] in "reports" and "descriptions" of the pictures "Western Harvest," "Labrador Fiord," and "Mountain River": (*a*) the frequency of mention, (*b*) the mean order of mention.

Each group was divided into random halves and coefficients of correlation were calculated between mean frequencies of each item for each half, between mean ranks, and also between mean frequency and mean rank for the whole group. The mean order for objects mentioned fewer than five times does not enter into these correlations.

Table I shows the coefficients calculated between mean frequencies obtained from reports.

---

[1] Subject to the limitation referred to above.

TABLE I.—COEFFICIENTS OF CORRELATION BETWEEN MEAN FREQUENCIES OBTAINED FROM REPORTS

Picture	Total No. of S's	No. of items	r
Mountain River	48	15	.75 ± .07
Mountain River	88	19	.86 ± .04
Labrador Fiord	55	25	.93 ± .03

It will be noted that a marked tendency exists for different groups of subjects to record items with the same relative frequency.

TABLE II.—COEFFICIENTS OF CORRELATION BETWEEN MEAN FREQUENCIES OBTAINED FROM DESCRIPTIONS

Picture	Total No. of S's	No. of items	r
Mountain River	41	22	.85 ± .04
Western Harvest	41	19	.95 ± .02

These figures exhibit the same tendency noted in the case of reports. That this relative frequency is substantially the same for "report" and "description" is evident from Table III.

TABLE III.—COEFFICIENTS OF CORRELATION BETWEEN MEAN FREQUENCIES OBTAINED FROM REPORTS AND DESCRIPTIONS

Picture	Total No. of S's	No. of items	r
Mountain River	82	22	.98 ± .01
Western Harvest	82	28	.98 ± .01

Tables IV, V and VI contain similar data in connection with the mean order of items observed.

TABLE IV.—COEFFICIENTS OF CORRELATION BETWEEN MEAN RANKS OBTAINED FROM REPORTS

Picture	Total No. of S's	No. of items	r
Mountain River	40	13	.77 ± .08
Mountain River	48	15	.76 ± .07
Mountain River	88	16	.64 ± .10
Western Harvest	40	13	.73 ± .09
Labrador Fiord	55	15	.62 ± .11

TABLE V.—COEFFICIENTS OF CORRELATION BETWEEN MEAN RANKS OBTAINED FROM DESCRIPTIONS

Picture	Total No. of S's	No. of items	r
Mountain River	41	15	.71 ± .09
Western Harvest	41	14	.72 ± .08

TABLE VI.—COEFFICIENTS OF CORRELATION BETWEEN MEAN RANKS OBTAINED FROM REPORTS AND DESCRIPTIONS

Picture	Total No. of S's	No. of items	r
Mountain River	80	18	.83 ± .05
Western Harvest	80	19	.83 ± .06

To obtain an idea to what extent *individual* subjects showed tendency to maintain the same order in report and description the following procedure was followed. From the alphabetized list of reagents nine were selected, *viz.*: First, sixth, eleventh, etc. For each of these the order in which each object was listed in report and description was tabulated and ρ calculated. These ρ's were converted into Pearson r's. Items named in only one of the two lists were treated as non-existent.

TABLE VII.—COEFFICIENTS OF CORRELATION BETWEEN ORDER OF ITEMS IN REPORT AND DESCRIPTION OF INDIVIDUAL SUBJECTS

No. of Reagent	r
1	.36
6	.78
11	.81
16	1.00
21	.87
26	.75
31	.69
36	.30
41	.67
Mean	.69

It is clear that a tendency exists towards employment of the same order of mention not only by groups but also by individuals. There is slight evidence that this tendency is not so pronounced as that to maintain the same relative frequencies.

Tables VIII and IX contain data bearing on the relationship between mean rank and mean frequency.

TABLE VIII.—COEFFICIENTS OF CORRELATION BETWEEN MEAN RANKS AND MEAN FREQUENCIES OBTAINED FROM REPORTS

Picture	Total No. of S's	No. of items	r
Mountain River	96	19	$-.49 \pm .10$
Mountain River	80	23	$-.66 \pm .09$
Western Harvest	80	19	$-.43 \pm .13$
Labrador Fiord	54	16	$-.76 \pm .07$
Labrador Fiord	56	16	$-.64 \pm .10$

TABLE IX.—COEFFICIENTS OF CORRELATION BETWEEN MEAN ORDERS AND MEAN FREQUENCIES OBTAINED FROM DESCRIPTIONS

Picture	Total No. of S's	No. of items	r
Mountain River	80	22	$-.64 \pm .09$
Western Harvest	80	19	$-.58 \pm .10$

The coefficients reported in Tables VIII and IX indicate the existence of a tendency for the more frequently "reported" or "described" items to be listed first and the less frequently, last.

This tendency was noted by Brown[1] in the lists of advertisements remembered in a given time by a group of students.

CONCLUSIONS

1. There is a fairly pronounced tendency for the items listed by groups in both "reports" and "descriptions" of pictures to be recorded with the same relative mean frequencies.

2. A fairly pronounced tendency exists for the items in "reports" and "descriptions" to be recorded in the same mean order.

3. Although considerable variation exists, the tendency noted in (2) alone is also true for individuals.

4. A tendency, not so clearly marked as those noted in (1) and (2) above, exists for the more frequently mentioned items to be recorded first and the less frequently last.

[1] Incidental Memory in a Group of Persons. *Psychological Review*, 1915.

THE DEVELOPMENT OF RESEARCH IN LEARNING

ROBERT A. DAVIS AND C. R. BALLARD

College of Education, University of Colorado

For some time college teachers have been apprehensive about the content of courses in educational psychology which have included the fields of measurement, statistics, child psychology, and learning. The majority of universities offer specific courses for treating the subject-matter listed under the general term educational psychology; consequently students taking this subject often duplicate previous work. The topic of learning frequently receives too little attention in educational psychology and this tendency is observed whether it is presented in a separate course or as one phase of a general course in educational psychology. This condition may be partially due to the inadequacy of research in learning and the fact that the available experiments have used types of materials which bear little relation to learning applicable in the schoolroom. Objective studies emphasize nonsense syllables, sorting of cards, and measure rote memory and motor skills without much attention to those types which are characteristic of learning in the classroom. It appears that research in learning is ill-adapted to classroom use, and that results of investigation are too conflicting to be of much value to the teacher of educational psychology. An analysis of objective studies in learning yields further evidence that this is a valid criticism.

Although there are obvious limitations to learning in its present stage of development the major emphasis in a course of educational psychology should be placed on learning and those factors which influence its economical development. This is particularly true of courses in educational psychology offered to students of education. Proceeding on the assumption that educational psychology should deal with learning and those factors which most directly affect it, an attempt was made to gather materials of a scientific nature which could be used in a course in advanced educational psychology designed for graduate students in education. Inasmuch as a psychological laboratory was not available, the class was conducted by means of discussion of experimental literature, each student organizing a paper on some subject with the scientific data which had been collected.

The character and amount of materials on learning and associated topics are presented here from articles and monographs found

in sixty-five periodicals[1] during a period of forty years. A majority of the studies were articles, although a considerable number were reported in monograph form. The term "studies" was used to include both types. The term "scientific" was used, as only studies of an experimental or statistical nature were considered. A majority of the investigations analyzed used the usual techniques of experimental research. The statistical investigations included those in which various types of statistical devices were employed, but where there was little attempt to control the findings by carefully controlled experimental procedures. Each study was examined for both content and method of research. Papers of a discursive or opinionated character were not included, nor were scientific investigations of a psychological character, although they obviously furnished a valuable background in the study of learning.

The studies collected were classified under fourteen major topics as indicated in Table I. The method employed was that of determining the number of studies and the number of pages of research produced according to decennial periods from 1890 to 1929 inclusive—a period of forty years. By classifying the materials in this way, it was believed that some interesting facts might be discovered regarding the changing tendencies in learning and associated topics for the period studied. It should be observed that the materials pertain to the general learning situation and do not show research for learning in specific school subjects. For the most part, the research has dealt with human learning, although there were included in some topics many experiments dealing with animal learning.

The results given in Table I show the distribution of studies according to decennial periods. They are listed according to the number and percentages of studies and pages allotted to each. Since the number of studies examined varied with the year periods, the important facts are found in the percentage columns which are computed by periods for studies and pages respectively. The most noticeable feature of this table is the marked increase in the number of studies since 1890. Attention is also directed to the fluctuation of various topics. Under Topic 1 (the nature and principles of learning) were included all those studies which dealt with the derivation or confirmation of theories, and the laws of learning. The data show that there were no studies during the first period, a perceptible increase during the third period and a

[1] These periodicals are listed at the end of the article.

TABLE I.—DISTRIBUTION OF LEARNING SUBJECT-MATTER ACCORDING TO DECENNIAL PERIODS

Topics	1890–1899				1900–1909				1910–1919				1920–1929				Studies, total		Pages, total	
	Studies		Pages		Studies		Pages		Studies		Pages		Studies		Pages					
	No.	Per cent	No.	Per cent	No.	Per cent	No.	Per cent	No.	Per cent	No.	Per cent	No.	Per cent	No.	Per cent	No.	Per cent	No.	Per cent
1. Nature and principles of learning	0	0	0	0	3	2.94	87	3.09	18	4.66	360	3.40	48	8.45	943	7.17	69	6.32	1,390	5.03
2. Relation of physical and mental ability	6	16.67	136	12.96	15	14.70	240	8.53	35	9.07	839	7.92	57	10.03	1,164	8.85	113	10.35	2,379	8.62
3. Observation, perception and report	0	0	0	0	5	4.90	126	4.48	6	1.16	95	.89	15	2.64	171	1.30	26	2.38	392	1.42
4. Reception of stimuli and mode of presentation	3	8.33	104	9.91	7	6.86	177	6.29	17	4.41	271	2.56	18	3.17	153	1.16	45	4.12	705	2.55
5. Motor learning	7	19.44	246	23.45	6	5.88	142	5.50	33	8.55	977	9.22	36	6.34	963	7.32	82	7.51	2,328	8.43
6. Mental and ideational learning	1	2.78	27	2.57	11	10.78	289	10.27	56	14.51	1,283	12.10	55	9.68	1,193	9.07	123	11.26	2,792	10.11
7. Permanence in learning	12	33.33	195	18.60	21	20.59	788	28.	68	17.62	1,833	17.29	79	13.91	1,780	13.54	180	16.48	4,596	16.65
8. Transference and interference in learning	2	5.55	25	2.38	15	14.70	378	13.43	50	12.95	2,087	19.69	39	6.87	1,189	9.04	106	9.71	3,679	13.32
9. Attention, interest and effort	1	2.78	27	2.57	3	2.94	84	2.98	21	5.44	588	5.55	28	4.93	317	2.41	53	4.85	1,016	3.68
10. Attitudes, incentives and motivation	0	0	0	0	2	1.96	37	1.31	18	4.66	316	2.98	77	13.56	1,728	13.14	97	8.88	2,081	7.54
11. Reasoning, problem solving, reflective thinking	1	2.78	63	6.01	2	1.96	55	1.95	10	2.59	315	2.97	16	2.82	569	4.33	29	2.66	1,002	3.63
12. Continuous work, fatigue, narcotics and drugs	1	2.78	32	3.05	9	8.82	334	11.87	43	11.14	1,207	11.39	49	8.63	1,058	8.05	102	9.34	2,631	9.53
13. Guidance in learning	1	2.78	109	10.39	1	.98	9	.03	4	1.04	70	.66	14	2.46	623	4.74	20	1.83	811	2.94
14. Drives and confidence	1	2.78	85	8.10	2	1.96	68	2.41	7	1.81	355	3.35	37	6.51	1,300	9.88	47	4.31	1,808	6.55
Totals	36	100.00	1,049	99.99	102	99.97	2,814	100.14	386	99.79	10,596	99.97	568	100.00	13,151	100.00	1,092	100.00	27,610	100.00

rapid rise in the fourth. This increase in the fourth period indicates during recent years a definite trend for psychologists to test experimentally the theories of learning and to check the commonly accepted laws of learning. As there has been a tendency to question the efficacy of the traditional laws of learning, a large number of theories of learning are now proposed, and it is only natural that these theories should be subjected to the experimental approach.

Topic 2 (the relation of physical and mental ability) included those studies which dealt with the relation of physical and mental traits in general using the various kinds of anthropometric measures, anatomical, dental and other ages together with the many studies on the relation of physical defects to intelligence and achievement. The greatest percentage of studies was found in the first two periods followed by a marked decline in the others. The early investigations were largely fact-finding and statistical studies which dealt, in the main, with the association of physical defects and teachers' estimates of ability and school marks. The later investigations include the use of new measures of maturity, and are characterized in their approach by greater objectivity. In recent years there have been only a few objective studies dealing with the relation of physical defects to intelligence, achievement and school progress. The slight rise in the fourth period over the preceding one indicates that this topic will be given renewed emphasis during the next ten years.

Under Topic 3 (observation, perception, and report) are those investigations which include the range and accuracy of observation, perception and those factors which influence its efficiency, the psychology of testimony and report ability. The facts show that there were no studies during the first period, but a relatively great amount during the second and a decline in the third with an increase in the fourth. In this country, Whipple has been one of the leaders in report both in interpreting the work of German scholars and in pointing out educational implications. His work and the more recent investigations of McGeoch have constituted the major portion of research on this subject. An analysis of this topic illustrates a tendency which exists among all those considered. Emphasis on research in certain topics fluctuates from period to period. There will be a number of studies for one period and one topic, and perhaps there will scarcely be any for another ten or twenty years; then there will be a renewed emphasis. At the present time there is a revived interest in observation, perception and report as is indicated by Fig. 1.

Topic 4 (the reception of stimuli and modes of presentation) covered those investigations which dealt with the question of ideational types, the efficiency of different modes of presentation in relation to learning, effectiveness of various forms of visual aids, and graphic tabular and textual methods of presenting quantitative data. This topic had its greatest emphasis during the first two periods with a

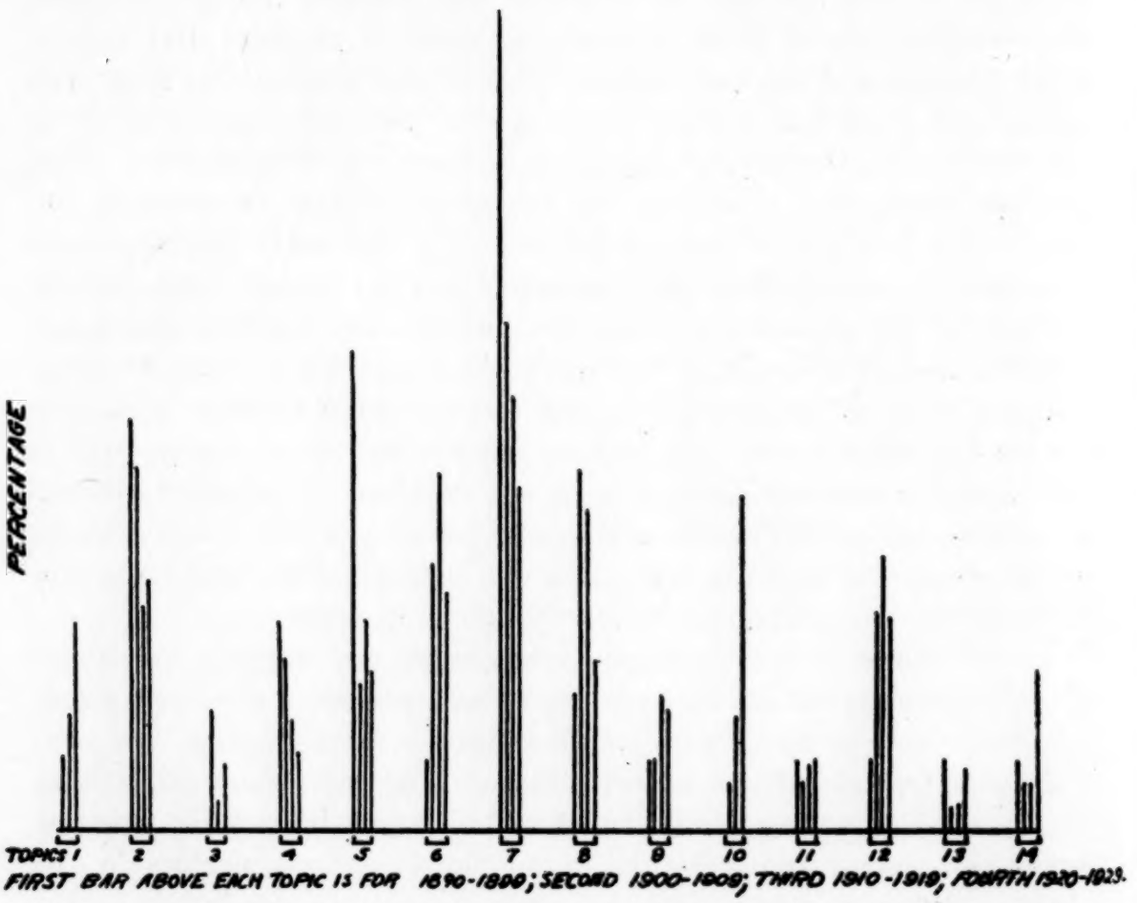

PERCENTAGE
TOPICS 1 2 3 4 5 6 7 8 9 10 11 12 13 14
FIRST BAR ABOVE EACH TOPIC IS FOR 1890-1899; SECOND 1900-1909; THIRD 1910-1919; FOURTH 1920-1929.

FIG. 1.—The distribution of learning subject matter according to decennial periods.

marked decline for the last two. The relatively great amount of research during the first two periods may be partially explained by the interest in the question of ideational types and the controversial issues involved. Although the discussion of types has been eliminated, there are a number of recent investigations which have been designed to test the validity of numerous modes of presentation as they apply to different school subjects and age-groups. Results of these investigations

indicate that there is still much conflict regarding the efficacy of the several modes of presentation, and that it is difficult to say whether the auditory, visual, or motor type of presentation is best. This field has hardly been opened, and considering the far-reaching importance of methods of presentation to learning, it is difficult to explain why there has been so little research produced on this subject.

Although the materials of motor and mental learning are ill-adapted to courses in learning and educational psychology, the investigations on these topics have been more extensive than others. The criticism has been that the investigations deal in the main with sensori-motor skills, nonsense syllables, etc., and that the results obtained from their use were not applicable to learning as it is found in the classroom. This criticism is found valid by analysis of the studies made. Under Topic 5 (motor learning) were included those investigations which dealt for instance with the tossing of balls, sorting of cards, walking tight wires, learning to typewrite and to play a piano. The term "sensory motor" is sometimes used to cover all the types of materials and functions under this heading. The results of the survey indicate that the greatest percentage of studies was made during the first period with a marked decline in the last two. The decrease in the last two periods may be attributed to interest in another type of learning—mental and ideational—which rivaled this topic during the last two periods. It is also to be expected that the skill and motor types of learning would receive their greatest emphasis during the early years as they are more adaptable to research methods than the less tangible materials classified as mental and ideational. Under mental and ideational learning were considered those studies which dealt with nonsense syllables, rational problems, sensible materials, mental work, etc. Although it is impossible to make a clear-cut distinction between the studies of a motor and mental character, an attempt was made to carefully distinguish between them. In motor learning, the individual reacts to a stimulus which is concrete in character, while in mental or ideational learning the individual is responding to stimuli largely through the medium of language. Language becomes a symbol which represents the stimulus in question. The two types of learning are often considered together, but it appears that there were sufficient inherent differences between the two types to justify a separate classification.

The largest percentage of studies is found in the second and third periods with a marked decrease in the fourth period. It would seem

that research on this topic has reached a saturation point. However, during the fourth period there has been some attempt to study a newer type of learning. Peterson is a pioneer in his attempt to make a rational learning test. Kuo, Lashley, and Muenzinger have also taken issue with the mechanistic viewpoint in learning and have proposed a psychological interpretation which accounts for learning in terms of purpose, insight, and plasticity. The investigations using nonsense materials and stressing rote learning have yielded little which can be of practical value in school. It is evident that in the future major emphasis will be given to learning of meaningful material and the rational process. However, research in this field will meet many obstacles because of the intangible nature of the functions measured and the difficulty of controlling experimental conditions.

Under Topic 7 (the permanence of learning) were considered those investigations which dealt with ways of measuring memory, types of materials, forgetting, age, sex, and intelligence. The aim was to consider the permanence of learning in its relation to all those factors which may influence it. Again, it is difficult to differentiate between permanence of learning and motor and ideational learning because the major method of measuring the efficiency of learning under any situation is through either immediate or remote memory. Although the largest percentage of studies is found during the first period, it is interesting to note that the other periods contain a relatively large percentage of studies.

Ranking third in importance for the forty-year period are the studies dealing with the transference and interference of learning. The transfer experiments are concerned with both transference in mental functions and school subjects. The studies bearing on interference were few, and those which have been reported are of recent origin. The largest percentage for this topic is found during the years from 1900 to 1920, with a marked decline from 1920 to 1929. However, the research on this topic has been characterized in recent years by improved experimental techniques in the form of control groups, and better equating and measurement making the results more convincing. The investigations on interference are suggestive and indicate that there is much that can be done in the future to clarify the problem of interference in learning. It would seem that this subject is of as much importance as the positive aspect of transference in learning.

The studies in Topic 9 (attention, interest and effort) show a marked increase from 1910 to 1929. These investigations have considered such problems as ways of measuring attention, rules governing acquirement of interest and the relation of attention and interest to effort. In recent years the investigations under this topic have been produced with the view of their application to industry, advertising and vocational psychology.

Research in Topic 10 (attitudes, incentives and motivation) shows one of the most outstanding tendencies of all the topics considered. Investigation in this field was hardly known until the fourth period, when there was a marked rise. The studies involve the measurement and significance of attitudes in learning, types of incentives in their relation to age, sex and intelligence, and the general problem of motivation. This important topic has been almost overlooked by writers of educational psychology in the past, but there is now available a large amount of valuable material.

The amount of research produced on Topic 11 (reasoning, problem solving and reflective thinking) has been seriously neglected throughout the entire forty-year period. Recent investigations have included the construction and standardization of scales for scientific thinking, the relation of reasoning to habit formation, thinking silently as compared with thinking aloud, and with the problem of the effect of training in reasoning on the ability to solve practical problems. This topic is usually given some attention in learning experiments of the ideational type, but, so far, the rational learning tests constitute about the only research in that field. This is a fertile topic for further investigation.

Topic 12 (continuous work, fatigue, drugs and narcotics) embraced those studies which dealt with the influence of fatigue, drugs and narcotics on mental and motor efficiency. The greatest percentage of research was produced during the second and third periods with a gradual decline in the fourth. Educational psychologists have usually given considerable space in textbooks to this problem although conclusions have not been convincing. Recent investigations have developed and elaborated many new techniques but the results are not far different from those of ten years ago.

The other topics listed—drives and confidence and guidance in learning—have rarely been considered through the forty-year period, but there is a new interest in these topics in recent years. Almost all studies on guidance have been made on animals; available results

suggest new methods of study and open a subject of especial significance to learning. Drives and confidence have received some attention in recent years, but with the exception of a few, all studies have been made with animals.

In order to throw further light on the data presented in Table I, the topics have been arranged in order of their importance as deter-

TABLE II.—TOPICS RANKED IN ORDER OF IMPORTANCE ACCORDING TO DECENNIAL PERIODS

Topic	1890–1899	1900–1909	1910–1920	1920–1929
	Rank[1]	Rank[2]	Rank[3]	Rank
1. Nature and principles of learning	12	9	8	6
2. Relation of physical and mental ability	3	2	5	3
3. Observation, perception and report	13	8	13	13
4. Reception of stimuli and method of presentation	4	6	10	11
5. Motor learning	2	7	6	9
6. Mental and ideational learning	6	4	2	4
7. Permanence in learning	1	1	1	1
8. Transference and interference in learning	5	3	3	7
9. Attention, interest, and effort	7	10	7	10
10. Attitudes, incentives and motivation	14	11	9	2
11. Reasoning, problem-solving, reflective thinking	8	12	11	12
12. Continuous work, fatigue, narcotics and drugs	9	5	4	5
13. Guidance in learning	10	14	14	14
14. Drives and confidence	11	13	12	8

[1] Topics 6, 9, 11, 12, 13, 14 of equal importance; also 1, 3, and 10.
[2] Topics 2 and 8 of equal importance; 1 and 9; 10, 11, and 14.
[3] Topics 1 and 10 are of equal importance.

mined by the percentages of the total studies for each devoted to the various topics. Since it was found that the relation between the percentage of studies and the percentage of pages was very close, the ranking below is based on studies alone. Those topics most important according to the percentage of studies come first, the next highest second and so on through all the topics.

Conclusions

The purpose of this paper was to analyze studies reported in periodicals and monographs with a view to noting trends, and to determine the lines along which research in learning is now being formulated. Some of the general conclusions may be outlined as follows:

1. The growth in research has been phenomenal during the past forty years. The most marked expansion has been during the later rather than the earlier periods. In recent years there has been a rapid development and expansion of new topics.

2. The study has indicated that the greatest amount of research for the entire forty-year period has dealt with the permanence of learning. Ranking next in importance according to the amount of research for the whole period is mental and ideational learning which is followed by the relation of physical and mental ability, transference and interference in learning, continuous work, fatigue, drugs and narcotics in the order named. The other topics have varying degrees of significance as has been indicated.

3. One of the most pronounced needs in research at present is on rational learning of the type which is emphasized in the classroom situation. There is also need for further development of the topics of reasoning, problem solving and reflective thinking; modes of presentation in relation to efficiency of learning; guidance in the learning situation; and observation, perception and report.

4. The study points out the need for surveys of this type at intervals of every few years. Such investigations would aid in revealing weaknesses in the various phases of the learning process, suggest the need for further research, yield valuable background to teachers and research workers, and stimulate students to carry forward timely projects in the field.

BOOK REVIEWS

Mental Measurement of Preschool Children, by Rachel Stutsman. Yonkers-on-Hudson: The World Book Company, 1930. Pp. X + 368.

The measurement of the mental capacities of the young child offers a unique challenge which has been met with marked success by Dr. Stutsman. Her book marks the culmination of extensive research during which a wide variety of materials have been experimented with in an effort to select those with most validity for measuring the mentality of preschool age children. Complete directions are given for the administration and scoring of her series of tests which is commonly known as the Merrill Palmer Scale. One of the most valuable features of the book is the inclusion of a guide for observing the child's social and emotional responses as he works with the test. The use of this scale enables the examiner to answer the question so often asked, "How did the child respond while he was taking the test?" "What was his attitude toward the test?" Such observations are of paramount importance in the examination of young children because they supplement the objective examination with developmental data and assist the examiner in evaluating objective results.

Of the tests retained in the final form of the scale some are language tests, some are tests of motor coordination and others are form board and picture tests. The criteria for the selection of materials were: Inherent interest for the child, utilization of a wide variety of abilities and activities, variation in difficulty of items, avoidance of materials influenced specifically by training and environment, simplicity and administration and scoring, and statistically validity. All the materials retained in the final revision of the test have been adequately standardized. Special studies of the influence of environment, sex and resistance on the test score have been made, the results of which are reported here. A comprehensive description of earlier efforts to construct mental tests for young children, most of which have been developed during the past ten years, is given. The only noticeable omissions are accounts of Russian experiments with psycho-motor tests of young children. Dr. Stutsman's book is recommended to all who wish the fullest account of the construction and administration of the Merrill Palmer Scale, as well as to those who wish to familiarize

themselves with the principles of testing young children and observing their behavior.

GERTRUDE HILDRETH.
Lincoln School of Teachers College, Columbia University.

The Psychology of the Elementary School, by Harry Grove Wheat. New York: Silver-Burdett and Co., 1931. Pp. VIII + 440.

Teachers need all the understanding of the psychology of elementary subjects which they can get. The teaching process is often carried on without an adequate appreciation of the nature of the learning process. Wheat's volume, therefore, is a timely one in that it gives us an additional text for courses in the psychology of elementary subjects to supplement the ones which are previously available.

With the exception of the first two and the last chapters the entire book is given to the psychology of particular subjects. Chapter I is on learning and teaching, Chapter II on the psychology of the school program, and Chapter XI on the psychology of the pupils' behavior. The chapters between these deal with the psychology of specific school subjects.

The book is somewhat in contrast with a recent one by Reed covering the same field, in that it is not primarily a summary of objective investigation and statistical studies. The author writes on in his own words, following his own outline, and conveying a message of his own, uninterrupted by numerous quotations of the findings of research workers. His statements and recommendations do not lack objective support, and there are numerous supporting footnotes but his plan of welding the entire mass of research material into a running account has made for readability and interest. This fact in particular will appeal to the undergraduate student who uses the volume as a text in a teacher-training course.

The boundary line between method and psychology is not always clear. The author has wisely refrained from making a sharp distinction on this point. At times he is pointing out the practical applications of a psychological principle; at other times he seems rather to be digging up a psychological reason for a method that is being advocated. If psychology has any excuse for existence, that excuse lies in its contribution to practice, and the author is to be congratulated upon his willingness to follow a psychological principle through to its application to teaching. This may cause some difficulty due to

his overlapping of methods courses, but a reasonable amount of overlapping is justifiable when the fundamental viewpoints of the courses are different.

The reviewer believes the volume will render a valuable service in teacher-training. C. C. CRAWFORD.

University of Southern California.

ROSCOE PULLIAM. *Extra-instructional Activities of the Teacher.* Pp. VI + 459. Garden City, New York: Doubleday, Doran and Company, Inc., 1930.

The pronounced increase in the "extra-instructional activities of teachers" creates a definite demand for a treatise on the subject. Although student-teachers present credit in general and in special methods of teaching, such courses are necessarily confined in the main to methods of creating learning situations. The beginning teacher needs additional help in meeting the immediate problems of the classroom, such as establishing classroom routine, maintaining good discipline, keeping accurate records, directing extra-curricular activities, etc. Furthermore, the teacher's relations to the community, to the social agencies, to fellow-teachers, and to the profession are likely to be as important as the ability to control classroom conditions. Pulliam's new book, "Extra-instructional Activities of the Teacher," is devoted to these two types of adjustments forced upon the novitiate.

Pulliam is an advocate of common sense in "school-keeping." Although many helps are listed, rigid school-routine and discipline practices are regarded as mere abuses, providing unnecessary occasion for conflict between teacher and pupil. In the modern school pupils come and go out and move about their rooms in the quickest, most reasonable order. Modern school discipline is inherent in the work of the school. Nevertheless, Pulliam has no emotional fervor for freedom. Freedom is an ideal to be approximated in practice.

Pulliam gives an excellent historical sketch of student participation in government, beginning with Lancaster and Bell and ending with French and McAndrew. The place and function of the school republic, the student council, the home-room, the assembly, and special projects are discussed in some detail. The dangers of an over-elaborate organization and of too much stress on police and monitorial duties are well stated.

Attendance and punctuality, the holding power of the school, and juvenile delinquency are logically grouped for discussion. Good teaching and worthwhile subject-matter are the "keys" to the solution of these related problems. Prizes and awards are condemned in no uncertain language. All three problems are related to the child's physical organism—physical defects, mentality, health—which is the subject for another unit. Although nothing new is here given, it is a good résumé of current literature on the topic.

In limited space, there is a good discussion of extra-curricular activities, especially of competitive school athletics. Says Pulliam: "When things have reached a point where an athletic coach often is more important in the public estimation than a principal, when a winning athletic team will completely cover all other pedogogical sins . . . when over-eager alumni . . . interfere with sound school policies . . . it is time to call a halt" (pp. 232–233). He then outlines a sound policy for school athletics (pp. 233–234). Other topics—publications, debating, literary societies, clubs, parties, assemblies,—receive the usual treatment.

The chapters treating relations with auxiliary educational agencies and with parents are especially noteworthy. Auxiliaries include public libraries, independent clubs—4H Clubs, Boy Scouts, Girl Scouts, Camp Fire Girls, Junior Red Cross, etc.—and parent-teachers associations. Especially valuable are the suggestions relating to patron's day. School exhibits are given a questionable status. Entertainments are viewed as excellent devices for school publicity. Customary commencement exercises are severely indicted. New-type report cards are presented for experimental purposes. The whole question of adequate records occupies another chapter.

The last section of the book is concerned with professional advancement and professional relations. Teaching could be but does not now constitute a profession. Devices for self-rating are advocated and discussed in some detail, with an excellent sample (pp. 383–390, inclusive). Training-in-service calls for the usual description of summer schools, extension courses, correspondence courses, reading circles, and teacher's meetings. Teacher participation in administration is advocated. The over-supervised teacher is given some hints. Supersensitive, combative, and jealous teachers are described in such a manner as to result in self-recognition. Especially interesting is the treatment of the teacher's freedom—much of the recent crusading by teachers is labelled "pure nonsense." The Illinois Federation

of Labor idea, that what the teacher does in private (or in public) is nobody's business, is called good in theory but not safe in practice. However, anyone with intelligence and character enough to teach is not likely to take liberties the public will not grant.

Although the book lacks originality, it is a good systematic treatment of those phases of classroom procedure, except actual methods, often neglected in "teacher-training" institutions, and rarely treated in any one book. The author is a superintendent of school who, no doubt, has recognized certain inadequacies in the newly employed teacher. This book consists of numerous, practical, empirical hints and helps in meeting classroom, community, and professional problems, and is designed to orient the inexperienced teacher and to stimulate the experienced. Because of the more or less encyclopedic style, it will find its greatest usefulness as required reading in professional courses—to supplement the usual methods text. Student-teachers might well read it while doing practice teaching—as part of the required work. It could be used in teachers' reading circles, although the topics invite much useless argument. All in all, it is a contribution to educational literature.

JOHN T. WAHLQUIST.
University of Utah.